

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:01:55 ; Search time 168 Seconds
(without alignments)
2079.198 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4392
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SSKOTDIPILSTQPMPEAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : Published Applications AA Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	5	US-10-781-564-1 Sequence 1, Appli
2	4385	99.8	836	3	US-09-899-569A-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-781-564-1
; Sequence 1, Application US/10781564
; Publication No. US20040247601A1
; GENERAL INFORMATION:
; APPLICANT: Quigley, James P.
; APPLICANT: Hooper, John D.
; APPLICANT: Teata, Jacqueline E.
; APPLICANT: The Scripps Research Institute
; TITLE OR INVENTION: Methods for Diagnosing Cancer and Decreasing Metastasis by Cancer
; FILE REFERENCE: 1361.0360US1
; CURRENT APPLICATION NUMBER: US/10/781,564
; PRIOR FILING DATE: 2004-02-18
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 836
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-781-564-1

Query Match 100.0%; Score 4392; DB 5; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGLNCGVSIALLGVLLGAARLRGAAAFIALPRESNITVLLKLGPTTLAKCYVI	60
DB	1	MAGLNCGVSIALLGVLLGAARLRGAAAFIALPRESNITVLLKLGPTTLAKCYVI	60
QY	61	SKRHITMISGERIVTFSCQSPENHFVLEIQKNIDCMGSCPFGVQLOPSTSLPT	120
DB	61	SKRHITMISGERIVTFSCQSPENHFVLEIQKNIDCMGSCPFGVQLOPSTSLPT	120
QY	121	LNRTFIMDVKAHKSIGLEQPSIRLROIGBESCPDGVTHSISGRIDATVVRIGTFCSN	180
DB	121	LNRTFIMDVKAHKSIGLEQPSIRLROIGBESCPDGVTHSISGRIDATVVRIGTFCSN	180
QY	181	GTISRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIISVFEGSGSATMSANY	240
DB	181	GTISRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIISVFEGSGSATMSANY	240
QY	241	PEGFPEDELMTQVVPVPAHLPAVSFLNFNLSNCKEKEVEYIIPGTTNPEVFKLBDK	300
DB	241	PEGFPEDELMTQVVPVPAHLPAVSFLNFNLSNCKEKEVEYIIPGTTNPEVFKLBDK	300
QY	301	QPGMAGNFNLSLQGCDDAQSFGILRLQFQVULVQHPONESNKIYVVDLSNERAMSLTIE	360
DB	301	QPGMAGNFNLSLQGCDDAQSFGILRLQFQVULVQHPONESNKIYVVDLSNERAMSLTIE	360
QY	361	PRPVKQSRKFPFGCFVCIESRTCSNLTITSGSHKISFLCDDTLRLMNVKTIISCTDH	420
DB	361	PRPVKQSRKFPFGCFVCIESRTCSNLTITSGSHKISFLCDDTLRLMNVKTIISCTDH	420
QY	421	RYCQRKSYSLQVPSDILHLPEVLEDFSWKLLVPRDRLSLVLPPOKLOQHHEKPCNTSF	480
DB	421	RYCQRKSYSLQVPSDILHLPEVLEDFSWKLLVPRDRLSLVLPPOKLOQHHEKPCNTSF	480
QY	481	SYLVASAIPSODLVFGSPCPGSGIKQIQOVKONISVTLRTFAPSFRQESRQGLTVSFLPY	540
DB	481	SYLVASAIPSODLVFGSPCPGSGIKQIQOVKONISVTLRTFAPSFRQESRQGLTVSFLPY	540
QY	541	FKERGVFTVPTDYSKYTLRTPNWDRLPSLITSVMNISVPRDOVACITFFKERSGVVQC	600
DB	541	FKERGVFTVPTDYSKYTLRTPNWDRLPSLITSVMNISVPRDOVACITFFKERSGVVQC	600
QY	601	TGRAFMIIQEBORTAERIPLSDEVDVLPKPSFHHHSFWNNTSNCSPTSQKQDLLFSVTLT	660
DB	601	TGRAFMIIQEBORTAERIPLSDEVDVLPKPSFHHHSFWNNTSNCSPTSQKQDLLFSVTLT	660
QY	661	PRTVDLTVILIAAGVGVLLLSALGLIICVKKKKKTKNGPAVINGNINTEMPQPK	720
DB	661	PRTVDLTVILIAAGVGVLLLSALGLIICVKKKKKTKNGPAVINGNINTEMPQPK	720
QY	721	KFQKGRKNDNSHVAVIEDTVMYGHLLQDSSGSLQPEVDVTRPFGQTMGVCPSPPTIC	780
DB	721	KFQKGRKNDNSHVAVIEDTVMYGHLLQDSSGSLQPEVDVTRPFGQTMGVCPSPPTIC	780
QY	781	SRAPTAKLATEEPPRSPSESESPYFHSNNNDVSSKOTDIPILSTQPMPEAE	836
DB	781	SRAPTAKLATEEPPRSPSESESPYFHSNNNDVSSKOTDIPILSTQPMPEAE	836

RESULT 2
US-09-899-569A-4
; Sequence 4, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweifler
; APPLICANT: Marwa Scherl-Mostaguer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher

; TITLE OF INVENTION: Tumorasoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 99.8%; Score 4385; DB 3; Length 836;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPGAAEAFETALPRESNITVLIKGPPTLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAARLPGAAEAFETALPRESNITVLIKGPPTLAKPCYIV 60
QY 61 SKRHITMLSIKGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLQSTSLPT 120
DB 61 SKRHITMLSIKGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLQSTSLPT 120
QY 121 LNRTIIMVKAHKSIGLEIQESIPLRQIGPESCPDGVTHSISGRIDATVIRIGTFCGN 180
DB 121 LNRTIIMVKAHKSIGLEIQESIPLRQIGPESCPDGVTHSISGRIDATVIRIGTFCGN 180
QY 181 GTVSRITKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRITKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDEIMTQFVVPALHRAVSFLNENLSNCERKEEVEYYIPGSTTNPDEVFKLEBK 300
DB 241 PEGFPEDEIMTQFVVPALHRAVSFLNENLSNCERKEEVEYYIPGSTTNPDEVFKLEBK 300
QY 301 OPGNAGNFNLSLOCDDOASPGILRLOFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 OPGNAGNFNLSLOCDDOASPGILRLOFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFVPGCFVCLERTCSSNLTITSGSKHKISFLCDDLTRLMMNVEKTI SCTDH 420
DB 361 PRPVQSRKFVPGCFVCLERTCSSNLTITSGSKHKISFLCDDLTRLMMNVEKTI SCTDH 420
QY 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVKDRLSLVLPAAQLQOHTHEKPCNTSF 480
DB 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVKDRLSLVLPAAQLQOHTHEKPCNTSF 480
QY 481 SYLVASATPSODLYFGSFCRGSSIKQIOYKONI SVTLRTFAPSFQOEAROGITVSPFIY 540
DB 481 SYLVASATPSODLYFGSFCRGSSIKQIOYKONI SVTLRTFAPSFQOEAROGITVSPFIY 540
QY 541 FAECEGVFTVPTPKSKVYLRTPNMWDGRLSTSVSNISVPRDOVACLTFPKEKSGVVCQ 600
DB 541 FAECEGVFTVPTPKSKVYLRTPNMWDGRLSTSVSNISVPRDOVACLTFPKEKSGVVCQ 600
QY 601 TGRAFMIIQEQRTAEIIFSLDEVL PKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
DB 601 TGRAFMIIQEQRTAEIIFSLDEVL PKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
QY 661 PRTVLTVLLIAVGGVLLSALGLITCVKKKKKTKNKGPVAGIYNGNINTEMPROPK 720
DB 661 PRTVLTVLLIAVGGVLLSALGLITCVKKKKKTKNKGPVAGIYNGNINTEMPROPK 720
QY 721 KFQGRKNDSHVYAVIEDTMYGHLLQDSSGSLQPEVDYTRPFGTGWGCPSPPTIC 780

DB 721 KFQGRKNDSHVYAVIEDTMYGHLLQDSSGSLQPEVDYTRPFGTGWGCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPPESESEPYTFSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPPESESEPYTFSHPNNGDVSSKOTDIPLLANTOEPMEPAE 836

Search completed: May 4, 2006, 21:05:31
Job time : 168 secs

QY 361 PRPVKQSKKFPVGCFCVCESTRCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
 DB 361 PRPVKQSKKFPVGCFCVCESTRCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
 QY 421 RYCQKSKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRSLVLVPAOKLOOHTHEKPCNTSF 480
 DB 421 RYCQKSKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRSLVLVPAOKLOOHTHEKPCNTSF 480
 QY 481 SYLVASAIPSODLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQZASRQGLTVSFIPY 540
 DB 481 SYLVASAIPSODLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQZASRQGLTVSFIPY 540
 QY 541 FKREGVFVTPDTKSKVYLRTPNMDRGPLSLTSVSMNISVPRDQVACLTFFKERSGVVCO 600
 DB 541 FKREGVFVTPDTKSKVYLRTPNMDRGPLSLTSVSMNISVPRDQVACLTFFKERSGVVCO 600
 QY 601 TGRAFMIIQEORTRAEELFSLDEVDLPKPSFPHHSFWNINISNCSPTSGKQDLFLFSVTLT 660
 DB 601 TGRAFMIIQEORTRAEELFSLDEVDLPKPSFPHHSFWNINISNCSPTSGKQDLFLFSVTLT 660
 QY 661 PRYVDLTVILAAVGGVLLLSALGLIICVKKKKKKTKNKPAGVINGNINTEMPPRPK 720
 DB 661 PRYVDLTVILAAVGGVLLLSALGLIICVKKKKKKTKNKPAGVINGNINTEMPPRPK 720
 QY 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDYTRPQGTMGVCPSPPTIC 780
 DB 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDYTRPQGTMGVCPSPPTIC 780
 QY 781 SRAPTAKLATEEPPRSPSESEPEYTFSSHNNGVSSKDTIDPLLSIOEPMEPAE 836
 DB 781 SRAPTAKLATEEPPRSPSESEPEYTFSSHNNGVSSKDTIDPLLSIOEPMEPAE 836

RESULT 2
 AEA01805
 ID AEA01805 standard; protein; 836 AA.
 AC AEA01805;
 XX
 XX 14-JUL-2005 (first entry)
 XX
 XX Human CUB domain containing protein 1 (CDCP1).
 DE
 XX CUB domain containing protein 1; CDCP1; ovary tumor; cytostatic;
 KW endocrine disease; genitourinary disease; gynecology and obstetrics;
 KW neoplasm; gene therapy; vaccine; diagnosis; drug screening.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 30..667
 FT /note="Predicted extracellular domain"
 XX
 XX MO2005042102-A1.
 XX
 XX 12-MAY-2005.
 XX
 XX 22-OCT-2004; 2004MO-GB004502.
 XX
 XX 22-OCT-2003; 2003GB-00024656.
 XX
 XX (CLUT) CELLTECH R & D LTD.
 XX
 XX Burgess NA;
 XX
 XX MPI; 2005-366521/37.
 DR N-PSDB; AEA01806.
 XX
 XX Use of CDCP1 polypeptide or an agent that interacts with or modulates the
 PT expression or activity of CDCP1 polypeptide for diagnosing, preventing or
 PT treating ovarian cancer or in screening for agents that may treat or
 PT prevent the disease.

XX
 PS Claim 6; SEQ ID NO 1; 48bp; English.
 CC The present sequence is the protein sequence of human CUB domain
 CC containing protein 1 (CDCP1). The invention is based on the finding that
 CC CDCP1 represents a novel therapeutic target for the treatment and/or
 CC prophylaxis of ovarian cancer. The invention relates to new uses of CDCP1
 CC in the diagnosis, screening, treatment and prophylaxis of ovarian cancer,
 CC and provides compositions comprising CDCP1 antibodies that are
 CC immunospecific for CDCP1, and agents that interact with or modulate the
 CC expression or activity of CDCP1 or which modulate the expression a
 CC nucleic acid encoding CDCP1. The CDCP1 polypeptide comprises the present
 CC sequence or is a derivative of this sequence. The use of an agent that
 CC interacts with or modulates the expression or activity of a CDCP1
 CC polypeptide for the manufacture of a medicament for the treatment and/or
 CC prophylaxis of ovarian cancer is claimed. The agent may be an antibody or
 CC its functionally active fragment, derivative or analog, which may be
 CC conjugated to a therapeutic molecule, detectable label, second antibody or
 CC fragment, effector or reporter molecule, cytotoxic agent or cytokine. A
 CC claimed method for the treatment and/or prophylaxis of ovarian cancer
 CC involves administering a composition comprising a CDCP1 polypeptide. The
 CC composition may be a vaccine. Also claimed are methods of using CDCP1 to
 CC screen for anti-ovarian cancer agents, and a method of screening for
 CC and/or diagnosis or prognosis of ovarian cancer in a subject, and/or
 CC monitoring the effectiveness of ovarian cancer therapy in a subject by
 CC detecting and/or quantifying expression of CDCP1 in a sample obtained
 CC from the subject.
 XX
 SQ Sequence 836 AA;

Query Match 100.0%; Score 4392; DB 9; Length 836;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGCVSIALGLYLIGARLPARGAFAETALPRESNTVILKGLPTTLARCYVI 60
 DB 1 MAGLNGCVSIALGLYLIGARLPARGAFAETALPRESNTVILKGLPTTLARCYVI 60
 QY 61 SKRHITMSTKSGEIVTFSCQSPENHFVEIOKNIDCMGSPCFGEVQLOPSTSLPT 120
 DB 61 SKRHITMSTKSGEIVTFSCQSPENHFVEIOKNIDCMGSPCFGEVQLOPSTSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLEQPSIRLRQIGBSCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLEQPSIRLRQIGBSCPDGVTHSISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRITKQEGVKAHLHPWHPRNVSQFSTIANSSIRLCITISVFGESATLMSANY 240
 DB 181 GTVSRITKQEGVKAHLHPWHPRNVSQFSTIANSSIRLCITISVFGESATLMSANY 240
 QY 241 PEGFPEDELMTMOPVVPARHLRAVSFLNPNLSNCRKEERYEYIIPGSTTPEVFKLBDK 300
 DB 241 PEGFPEDELMTMOPVVPARHLRAVSFLNPNLSNCRKEERYEYIIPGSTTPEVFKLBDK 300
 QY 301 QPGNAGNFNLSTQCDQDAQSPGILRLQFOVLVOHPQNESNKIYVVDLSNERAMSLTIE 360
 DB 301 QPGNAGNFNLSTQCDQDAQSPGILRLQFOVLVOHPQNESNKIYVVDLSNERAMSLTIE 360
 QY 361 PRPVKQSKKFPVGCFCVCESTRCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
 DB 361 PRPVKQSKKFPVGCFCVCESTRCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
 QY 421 RYCQKSKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRSLVLVPAOKLOOHTHEKPCNTSF 480
 DB 421 RYCQKSKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRSLVLVPAOKLOOHTHEKPCNTSF 480
 QY 481 SYLVASAIPSODLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQZASRQGLTVSFIPY 540
 DB 481 SYLVASAIPSODLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQZASRQGLTVSFIPY 540
 QY 541 FKREGVFVTPDTKSKVYLRTPNMDRGPLSLTSVSMNISVPRDQVACLTFFKERSGVVCO 600
 DB 541 FKREGVFVTPDTKSKVYLRTPNMDRGPLSLTSVSMNISVPRDQVACLTFFKERSGVVCO 600

QY 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPGSKOLDLFSVTLT 660
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 836 AA.
Query Match 99.8%, Score 4385, DB 5, Length 836,
Best Local Similarity 99.8%, Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPGSKOLDLFSVTLT 660
Db 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPGSKOLDLFSVTLT 660
QY 661 PRYVDLTVIIILAAVGGVLLLSALGLIICVKKKKKTKNGPAVGINGNINTEMPROPK 720
Db 661 PRYVDLTVIIILAAVGGVLLLSALGLIICVKKKKKTKNGPAVGINGNINTEMPROPK 720
QY 721 KFOGRKNDSSHVAIVIEDTMVYGHLLQDSSGSLQPEVDVTRPFGTMGVCPPSPPTIC 780
Db 721 KFOGRKNDSSHVAIVIEDTMVYGHLLQDSSGSLQPEVDVTRPFGTMGVCPPSPPTIC 780
QY 781 SRAPTAKLATEEPPRPPSPSESEPTTFSSHNNGVSSKOTDIPLLSTOEPMPEAE 836
Db 781 SRAPTAKLATEEPPRPPSPSESEPTTFSSHNNGVSSKOTDIPLLSTOEPMPEAE 836
RESULT 3
ABP69553
ID ABP69553 standard; protein; 836 AA.
XX
AC ABP69553;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polypeptide SEQ ID NO 1600.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytotoxic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antirheumatic.
XX
OS Homo sapiens.
XX
PN MO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US0005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AI, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; AB211770.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1600; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB21119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP6902-ABP6949) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,

CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 836 AA.
Query Match 99.8%, Score 4385, DB 5, Length 836,
Best Local Similarity 99.8%, Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGINCYSIALLVLLIGAAIRPRGAFAFIALPRESNTIVILKGTPTLLAKCYIV 60
Db 1 MAGINCYSIALLVLLIGAAIRPRGAFAFIALPRESNTIVILKGTPTLLAKCYIV 60
QY 61 SKRHITMLSTIKSGRIYFTFSCQSPENHFVIEIKNTIDMSGPCPGFVOLQPTSLPT 120
Db 61 SKRHITMLSTIKSGRIYFTFSCQSPENHFVIEIKNTIDMSGPCPGFVOLQPTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLELQFSIRLRQIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMDVKAHKSIGLELQFSIRLRQIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMGQGVKALHLPMFHPNVSGFSIANSSIKRLCTIISVFEBSATLMSANY 240
Db 181 GTVSRIKMGQGVKALHLPMFHPNVSGFSIANSSIKRLCTIISVFEBSATLMSANY 240
QY 241 PEGPPEELIMQOVVPAHILASVSLFNPNLSNERKEERYIIPGTTNPEVKEDK 300
Db 241 PEGPPEELIMQOVVPAHILASVSLFNPNLSNERKEERYIIPGTTNPEVKEDK 300
QY 301 QPGMAGNFMNLSOGCQDQASPGILRLQFQVILQHPONESNKIYVVDLSNERAMSLTIE 360
Db 301 QPGMAGNFMNLSOGCQDQASPGILRLQFQVILQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKSRKVPFGCVLCESRTCSNLTITSGSHKISFLCDDTLRLMNVEXTISCTDH 420
Db 361 PRPVKSRKVPFGCVLCESRTCSNLTITSGSHKISFLCDDTLRLMNVEXTISCTDH 420
QY 421 RYCQKSYSLQVPSDILHLPELHDPFWKLLVPDRSLIVPPOKLOQHTHEPCNTSF 480
Db 421 RYCQKSYSLQVPSDILHLPELHDPFWKLLVPDRSLIVPPOKLOQHTHEPCNTSF 480
QY 481 SYLVASAIPODLYFGSFCRGSIKQIOVKONISVTLRTFAFSROBASRQGLVSPFIPY 540
Db 481 SYLVASAIPODLYFGSFCRGSIKQIOVKONISVTLRTFAFSROBASRQGLVSPFIPY 540
QY 541 FKEGVFTVTPDTKSKYLLRTPNMDRGLPSLTSVSNISVPRDVACTFFKERSGVVCC 600
Db 541 FKEGVFTVTPDTKSKYLLRTPNMDRGLPSLTSVSNISVPRDVACTFFKERSGVVCC 600
QY 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPGSKOLDLFSVTLT 660
Db 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPGSKOLDLFSVTLT 660
QY 661 PRYVDLTVIIILAAVGGVLLLSALGLIICVKKKKKTKNGPAVGINGNINTEMPROPK 720
Db 661 PRYVDLTVIIILAAVGGVLLLSALGLIICVKKKKKTKNGPAVGINGNINTEMPROPK 720
QY 721 KFOGRKNDSSHVAIVIEDTMVYGHLLQDSSGSLQPEVDVTRPFGTMGVCPPSPPTIC 780
Db 721 KFOGRKNDSSHVAIVIEDTMVYGHLLQDSSGSLQPEVDVTRPFGTMGVCPPSPPTIC 780
QY 781 SRAPTAKLATEEPPRPPSPSESEPTTFSSHNNGVSSKOTDIPLLSTOEPMPEAE 836
Db 781 SRAPTAKLATEEPPRPPSPSESEPTTFSSHNNGVSSKOTDIPLLSTOEPMPEAE 836
RESULT 4
AAM49641
ID AAM49641 standard; protein; 836 AA.

XX AAM49641;
AC 17-MAY-2002 (first entry)
XX
XX Human tumour-associated antigen B345 protein SEQ ID NO 4.
DE
XX Tumour-associated antigen; human; B345; cytostatic; cell communication;
XX cell interaction; signal transduction; metastasis; cancer; colon;
KM immunotherapy; carcinoma; lung; diagnosis.
XX
XX Homo sapiens.
OS
XX MO200204508-A1.
XX
XX 17-JAN-2002.
PD
XX
XX 05-JUL-2001; 2001WO-EP007705.
PF
XX
XX 07-JUL-2000; 2000DE-01033080.
PR
XX 19-APR-2001; 2001DE-01019294.
PR
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
PA
PI Schweifer N, Scherl-Mostageer M, Sommergruber W, Abseher R;
XX
XX WPI: 2002-171704/22.
DR
XX N-PSDB; ABA99507.
PT
XX New tumor-associated antigen B345, useful for diagnosis and immunotherapy
PT of tumors, also related nucleic acid and antibodies.
XX
XX
PS Claim 1; Page 85-88; 102pp; German.
XX
XX This invention describes a novel tumour-associated antigen, designated
CC B345 which has cytostatic activity. B345 is involved in communication,
CC interaction and/or signal transduction with extracellular components and
CC ligands, especially in the metastatic potential of cancers, particularly
CC of the colon. B345 or its immunogenic fragments, also the DNA that
CC encodes it, are useful for immunotherapy of cancer, particularly
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for
CC treatment and diagnosis of cancers that are associated with B345
CC expression, including their use for targeted delivery of cytotoxic or
CC radioactive agents. Probes derived from B345 can be used to detect tumour
CC -specific mutations in the B345 sequence, and can be used to screen for
CC B345 specific modulators. This sequence represents a human B345 tumour-
CC associated antigen described in the invention
CC
XX
XX Sequence 836 AA;
SQ
Query Match 99.8%; Score 4385; DB 5; Length 836;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGLNCGVSIALLGVLLGAAALPRGAFAFEIALPRESNITVLILKIGTPTLLAKPCYIVI 60
DB 1 MAGLNCGVSIALLGVLLGAAALPRGAFAFEIALPRESNITVLILKIGTPTLLAKPCYIVI 60
QY 61 SKRHITMTLSIKSGERIVFTFSCSPENHFVIEIQKNIDMSGRCPPGEVQLQPSSTLLPT 120
DB 61 SKRHITMTLSIKSGERIVFTFSCSPENHFVIEIQKNIDMSGRCPPGEVQLQPSSTLLPT 120
QY 121 LNRTFTWVKAHKSIGLEIOPSIPLRLOIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFTWVKAHKSIGLEIOPSIPLRLOIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHLPMFHPRVNVGSFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKALHLPMFHPRVNVGSFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRLASVSLFNPLNSNCERKEERVYIIPGSTTNPDEVFKLBDK 300
DB 241 PEGFPEDELMTWQFVVPALRLASVSLFNPLNSNCERKEERVYIIPGSTTNPDEVFKLBDK 300

QY 301 QPGNMGFNLSLQCCDDAOSPGILRLQFOVLVQHPONEBNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNMGFNLSLQCCDDAOSPGILRLQFOVLVQHPONEBNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSRKFPVGGFCVLESRTCSNLTJSGSKHKSIFLCDLTRLMMNVKTIISCTDH 420
DB 361 PRPVKQSRKFPVGGFCVLESRTCSNLTJSGSKHKSIFLCDLTRLMMNVKTIISCTDH 420
QY 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLLVPPDRSLVVPQKLOQHHEKPCNVSF 480
DB 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLLVPPDRSLVVPQKLOQHHEKPCNVSF 480
QY 481 SYLVASAIPOSQDLFGSFCPEGSIKQIOVKONISVTLTRTFAPSPROEASROGLTVSFIY 540
DB 481 SYLVASAIPOSQDLFGSFCPEGSIKQIOVKONISVTLTRTFAPSPROEASROGLTVSFIY 540
QY 541 FKEGCVFTVTPDTSKYVLRTPNMDRGLPSLTSVSNISVPRDQVACLTFFKERSGVYQ 600
DB 541 FKEGCVFTVTPDTSKYVLRTPNMDRGLPSLTSVSNISVPRDQVACLTFFKERSGVYQ 600
QY 601 TGRAFMIIQEORTAEBEIFSLEDEVLPKPSFHHSFWVNIENCSPTSGKQDLLEFSVTLT 660
DB 601 TGRAFMIIQEORTAEBEIFSLEDEVLPKPSFHHSFWVNIENCSPTSGKQDLLEFSVTLT 660
QY 661 PRTVDLTVILAAVGGVLLSALGLIICVKKKKKKTKNKGPAGIYNGNINTEMPROPK 720
DB 661 PRTVDLTVILAAVGGVLLSALGLIICVKKKKKKTKNKGPAGIYNGNINTEMPROPK 720
QY 721 KFOGRKNDSDHYAVIEDTMYGHLLODSSGSFLQPEVDITYRPPQGTWGVCPSPPTIC 780
DB 721 KFOGRKNDSDHYAVIEDTMYGHLLODSSGSFLQPEVDITYRPPQGTWGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESPYTFSHNNNDVSKDIDILLSTQEMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESPYTFSHNNNDVSKDIDILLSTQEMEPAE 836

Search completed: May 4, 2006, 20:56:49
Job time : 190 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:02:50 ; Search time 28 Seconds

(without alignments)
1381.930 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGLNCGVSIALLGVLLIGA.....SSKDTDIPILSTQEPMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : Published Applications AA New*

1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /SIDS5/ptodata/2/pubpaa/PCF_NEW_PUB.pep.*
6: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
7: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
8: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
9: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
10: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
11: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
12: /SIDS5/ptodata/2/pubpaa/US66_NEW_PUB.pep1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	11	US-11-050-857-989
2	4385	99.8	836	11	US-11-050-857-988
3	4059	92.4	770	11	US-11-050-857-252

ALIGNMENTS

RESULT 1
US-11-050-857-989
; Sequence 989, Application us/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 989

LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-989

Query Match 100.0%; Score 4392; DB 11; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGLNCGVSIALLGVLLIGAARLPAGAAFEIALPRESNITVLIKLGPTTLAKPCYIV	60
DB	1	MAGLNCGVSIALLGVLLIGAARLPAGAAFEIALPRESNITVLIKLGPTTLAKPCYIV	60
QY	61	SKRHITMLSTKSGERIVTFECQSPENHFVIEIQNIDCMGSPCPRGVOLOPSTSLPT	120
DB	61	SKRHITMLSTKSGERIVTFECQSPENHFVIEIQNIDCMGSPCPRGVOLOPSTSLPT	120
QY	121	LNRTFIMDKAHKSIIGLELOFSIPRLRQIGGESCPCDVTHSISGRIDATVVRIGTFCSN	180
DB	121	LNRTFIMDKAHKSIIGLELOFSIPRLRQIGGESCPCDVTHSISGRIDATVVRIGTFCSN	180
QY	181	GTISRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGSGATLMSANY	240
DB	181	GTISRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGSGATLMSANY	240
QY	241	PEGPEDELMTWQFVVPALHRAVSFLNPNINSCKERKEERYEYTPGSTNPEVFKLBDK	300
DB	241	PEGPEDELMTWQFVVPALHRAVSFLNPNINSCKERKEERYEYTPGSTNPEVFKLBDK	300
QY	301	QPNMAGNFNLSLQGGDDAOSPGILRQFOVLVQHPONESNKIYVVDLSNERAMSLITE	360
DB	301	QPNMAGNFNLSLQGGDDAOSPGILRQFOVLVQHPONESNKIYVVDLSNERAMSLITE	360
QY	361	PRPVQSKKFPVPGCVCLSESTCSNLTLSGSKHKISFLDCLDLRLMANNEXTISCTDH	420
DB	361	PRPVQSKKFPVPGCVCLSESTCSNLTLSGSKHKISFLDCLDLRLMANNEXTISCTDH	420
QY	421	RYCORKSYSLQVPSDILHLPELHDFSWKLLVPKDRSLVLVPAQKLOQTHHEKPCNTSF	480
DB	421	RYCORKSYSLQVPSDILHLPELHDFSWKLLVPKDRSLVLVPAQKLOQTHHEKPCNTSF	480
QY	481	SYLVASAIPSQDLVFGSCPCGSGIKQIOVKONISVTLTFAPSFQESROGLTFSFIPY	540
DB	481	SYLVASAIPSQDLVFGSCPCGSGIKQIOVKONISVTLTFAPSFQESROGLTFSFIPY	540
QY	541	FKEEGVFTVPTDTSKYVLRTPNMDRGLPSLTSVWNISVPRDQVACLTFFKERSGVVCO	600
DB	541	FKEEGVFTVPTDTSKYVLRTPNMDRGLPSLTSVWNISVPRDQVACLTFFKERSGVVCO	600
QY	601	TGRAFMIIQEQRTAAEIEFSLDEDLVLPKPSFHHSFWVNISNCSPFSKQDLFLFSVLT	660
DB	601	TGRAFMIIQEQRTAAEIEFSLDEDLVLPKPSFHHSFWVNISNCSPFSKQDLFLFSVLT	660
QY	661	PRYDLYTILAAAGGVLLLSALGLTICCYKTKKKTKNKPAVGINVGNINTEMPBPCK	720
DB	661	PRYDLYTILAAAGGVLLLSALGLTICCYKTKKKTKNKPAVGINVGNINTEMPBPCK	720
QY	721	KFOKGRKNDSHVAVIVEDTWYGHLODSGSGFLQPEVDYTRPFGTMGVCPSPPTIC	780
DB	721	KFOKGRKNDSHVAVIVEDTWYGHLODSGSGFLQPEVDYTRPFGTMGVCPSPPTIC	780
QY	781	SRAPTAKLATEEPSPSESESEBYTFSHPNNGDVSSKDTDIPILSTQEPMEPAE 836	
DB	781	SRAPTAKLATEEPSPSESESEBYTFSHPNNGDVSSKDTDIPILSTQEPMEPAE 836	

RESULT 2
US-11-050-857-988
; Sequence 988, Application us/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS

; TITLE OF INVENTION: THEROF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 988
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-988

Query Match 99.8%; Score 4385; DB 11; Length 836;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAGLNCGVSIALLGVLLGGAARLPGAAEAFIALPRESNITVLIKLGPTLLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGGAARLPGAAEAFIALPRESNITVLIKLGPTLLAKPCYIYI 60
OY 61 SKRHITMISIKSGRIIVFTFSCSPENHVEIIOKNIDCMGCPCEPGEVQLOPSTSLPT 120
DB 61 SKRHITMISIKSGRIIVFTFSCSPENHVEIIOKNIDCMGCPCEPGEVQLOPSTSLPT 120
OY 121 LNRTFIMDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
OY 181 GTVSHIKQEGVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANY 240
DB 181 GTVSHIKQEGVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANY 240
OY 241 PEGFEDELMTWQFVVPALRLASVSFLNFNLSNCRKEEREVEYIIPGSTNPEVFKLEDK 300
DB 241 PEGFEDELMTWQFVVPALRLASVSFLNFNLSNCRKEEREVEYIIPGSTNPEVFKLEDK 300
OY 301 QGNAGNFNLSIQCCDDAQSPGILRLQFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QGNAGNFNLSIQCCDDAQSPGILRLQFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
OY 361 PRPVKSRKRVPGFCVLCESRTCSSNLITSGSKHISFLCDDLRLMANNVEKTIISCTDH 420
DB 361 PRPVKSRKRVPGFCVLCESRTCSSNLITSGSKHISFLCDDLRLMANNVEKTIISCTDH 420
OY 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSF 480
DB 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSF 480
OY 481 SYLVASAIIPSDILYFGSFCPGGSIKOIOVKONISVTLRTPAFSPROEASRQGLTVSFY 540
DB 481 SYLVASAIIPSDILYFGSFCPGGSIKOIOVKONISVTLRTPAFSPROEASRQGLTVSFY 540
OY 541 FKEBGFVTPDPYKSKVYLRTNMDRGRLSLTSVSNISVPRDQVACLTFPKERSGVVQ 600
DB 541 FKEBGFVTPDPYKSKVYLRTNMDRGRLSLTSVSNISVPRDQVACLTFPKERSGVVQ 600
OY 601 TGRAMITIOEQRTABEIFSLEDEVLPRKSPFHHSFWNINISCSPTSGQDLDFSVTLT 660
DB 601 TGRAMITIOEQRTABEIFSLEDEVLPRKSPFHHSFWNINISCSPTSGQDLDFSVTLT 660
OY 661 PRTVDLTVLLIAVGGVLLLSALGLIICVYKKKKKTKNGPAVGIVNGNINTEMPROK 720
DB 661 PRTVDLTVLLIAVGGVLLLSALGLIICVYKKKKKTKNGPAVGIVNGNINTEMPROK 720
OY 721 KFQKRKNDSHVAVIEDTMYGHLLQDSSGFLQPEVDYTRPFQGTMGVCPSPPTIC 780
DB 721 KFQKRKNDSHVAVIEDTMYGHLLQDSSGFLQPEVDYTRPFQGTMGVCPSPPTIC 780
OY 781 SAAPPAKALATEPRPSPSESESEPTTSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836
DB 781 SAAPPAKALATEPRPSPSESESEPTTSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836

RESULT 3

US-11-050-857-252
; Sequence 252, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 252
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-252

Query Match 92.4%; Score 4059; DB 11; Length 770;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 768; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 67 MUISKGRIVFTFSCSPENHVEIIOKNIDCMGCPCEPGEVQLOPSTSLPTLNRTFI 126
DB 1 MUISKGRIVFTFSCSPENHVEIIOKNIDCMGCPCEPGEVQLOPSTSLPTLNRTFI 60
OY 127 MDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSNIGVSR 186
DB 127 MDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSNIGVSR 120
OY 187 KMQEGVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANYEGEPPE 246
DB 187 KMQEGVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANYEGEPPE 180
OY 247 DELMTWQFVVPALRLASVSFLNFNLSNCRKEEREVEYIIPGSTNPEVFKLEDKOPGNA 306
DB 247 DELMTWQFVVPALRLASVSFLNFNLSNCRKEEREVEYIIPGSTNPEVFKLEDKOPGNA 240
OY 307 GNFNLSIQCCDDAQSPGILRLQFVLYOHPONESNKIYVVDLSNERAMSLTIEBRPVKQ 366
DB 307 GNFNLSIQCCDDAQSPGILRLQFVLYOHPONESNKIYVVDLSNERAMSLTIEBRPVKQ 300
OY 367 SRKFPVPGFCVLCESRTCSSNLITSGSKHISFLCDDLRLMANNVEKTIISCTDHRVCOR 426
DB 367 SRKFPVPGFCVLCESRTCSSNLITSGSKHISFLCDDLRLMANNVEKTIISCTDHRVCOR 360
OY 427 SYSLOVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSFYLVAS 486
DB 427 SYSLOVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSFYLVAS 420
OY 487 AIPSDILYFGSFCPGGSIKOIOVKONISVTLRTPAFSPROEASRQGLTVSFYKEG 546
DB 487 AIPSDILYFGSFCPGGSIKOIOVKONISVTLRTPAFSPROEASRQGLTVSFYKEG 480
OY 547 FTVTDPDTSKYVYLRTNMDRGRLSLTSVSNISVPRDQVACLTFPKERSGVVQCGTGRAFM 606
DB 547 FTVTDPDTSKYVYLRTNMDRGRLSLTSVSNISVPRDQVACLTFPKERSGVVQCGTGRAFM 540
OY 607 IIOEQRTABEIFSLEDEVLPRKSPFHHSFWNINISCSPTSGQDLDFSVTLT 666
DB 607 IIOEQRTABEIFSLEDEVLPRKSPFHHSFWNINISCSPTSGQDLDFSVTLT 600
OY 667 TVILLIAVGGVLLLSALGLIICVYKKKKKTKNGPAVGIVNGNINTEMPROK 726
DB 667 TVILLIAVGGVLLLSALGLIICVYKKKKKTKNGPAVGIVNGNINTEMPROK 660
OY 727 KNDSHVAVAVIEDTMYGHLLQDSSGFLQPEVDYTRPFQGTMGVCPSPPTICRAPTA 786
DB 727 KNDSHVAVAVIEDTMYGHLLQDSSGFLQPEVDYTRPFQGTMGVCPSPPTICRAPTA 720
OY 787 KLAATEPRPSPSESESEPTTSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836
DB 787 KLAATEPRPSPSESESEPTTSHPNNGDVSSKOTDIPLLSTOEPMEPAE 770

Fri May 5 16:14:27 2006

us-10-781-564-1.pct.rapbn

Page 3

Search completed: May 4, 2006, 21:06:06
Job time : 28 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 20:57:05 ; Search time 45 Seconds
(without alignments)

1787.494 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGINCGVSTALLGVLLGA.....SSKDTDIPLSTQEPMPAP 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description

No matches found					

Search completed: May 4, 2006, 21:01:42
Job time : 45 secs

THIS PAGE LEFT BLANK

THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:01:55 ; Search time 168 Seconds
(without alignments)
2079.198 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4392
Sequence: 1 MAGLNGCVSIALLGVLGAA.....SSKTDIPILSTQEPMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US108_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	5	US-10-781-564-1
2	4385	99.8	836	3	US-09-899-569A-4

ALIGNMENTS

RESULT 1
US-10-781-564-1
; Sequence 1, Application US/10781564
; Publication No. US20040247601A1
; GENERAL INFORMATION:
; APPLICANT: Quigley, James P.
; APPLICANT: Hooper, John D.
; APPLICANT: Teeta, Jacqueline E.
; TITLE OF INVENTION: Methods for Diagnosing Cancer and Decreasing Metastasis by Cancer
; FILE REFERENCE: 1361.036US1
; CURRENT APPLICATION NUMBER: US/10/781,564
; PRIOR APPLICATION NUMBER: 2004-02-18
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 836
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-781-564-1

Query Match 100.0%; Score 4392; DB 5; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGLNGCVSIALLGVLGAAARLPGAAEFAIALPRESNITVLKIGFTLLAKPCYIV	60
DB	1	MAGLNGCVSIALLGVLGAAARLPGAAEFAIALPRESNITVLKIGFTLLAKPCYIV	60
QY	61	SKRHITMISIKGERIVFTFSCQSPENHVFIIQKIDCMGSPCPGEGVQLPSTILPT	120
DB	61	SKRHITMISIKGERIVFTFSCQSPENHVFIIQKIDCMGSPCPGEGVQLPSTILPT	120
QY	121	LNRTFIWDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGIDATVIRIGFC	180
DB	121	LNRTFIWDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGIDATVIRIGFC	180
QY	181	GVSRIRKQEGVKALHLPWFPRNVSGFSIANRSIKRLCTIESVFEGSGATLMSANY	240
DB	181	GVSRIRKQEGVKALHLPWFPRNVSGFSIANRSIKRLCTIESVFEGSGATLMSANY	240
QY	241	PGFPPEDELMTWQFVVPALRASVSFLNPNLSNCRKERRVYIIPGTTNPEVFL	300
DB	241	PGFPPEDELMTWQFVVPALRASVSFLNPNLSNCRKERRVYIIPGTTNPEVFL	300
QY	301	OPGNMAGNPNLSLOGCDODDASPGILRLOFVLVQHPQESNKIYVVDLSNERAM	360
DB	301	OPGNMAGNPNLSLOGCDODDASPGILRLOFVLVQHPQESNKIYVVDLSNERAM	360
QY	361	PRPVQSRKFVGCVCVCESRCSNLTLSGSKKHSIFLCDDLTRLMANVETKISCTD	420
DB	361	PRPVQSRKFVGCVCVCESRCSNLTLSGSKKHSIFLCDDLTRLMANVETKISCTD	420
QY	421	RYCQKRSISLOVPSITLHLPELHDPMSKLVLPKORLSLVYPAQLOQHTHEKPC	480
DB	421	RYCQKRSISLOVPSITLHLPELHDPMSKLVLPKORLSLVYPAQLOQHTHEKPC	480
QY	481	RYCQKRSISLOVPSITLHLPELHDPMSKLVLPKORLSLVYPAQLOQHTHEKPC	480
DB	481	RYCQKRSISLOVPSITLHLPELHDPMSKLVLPKORLSLVYPAQLOQHTHEKPC	480
QY	541	FKBEGVFTVPTDPTKSKVLRTPMMDRGLPSLTSVSMNISVPRDOVACLTFFK	600
DB	541	FKBEGVFTVPTDPTKSKVLRTPMMDRGLPSLTSVSMNISVPRDOVACLTFFK	600
QY	601	TGRAMITIOEORTRAEIIFSLDEDEVLPKPSFHHHSFWVNISNCSPTSGKQD	660
DB	601	TGRAMITIOEORTRAEIIFSLDEDEVLPKPSFHHHSFWVNISNCSPTSGKQD	660
QY	661	PTVDLTVILIAAAGGVLLLSALGLIICVYKKKKKTKNGAVGIYNININTEM	720
DB	661	PTVDLTVILIAAAGGVLLLSALGLIICVYKKKKKTKNGAVGIYNININTEM	720
QY	721	KFOGRKNDKSHVYVIEDTWYGYGLLQDSGSLQPEVDYTRPPQGTGVCVPSPT	780
DB	721	KFOGRKNDKSHVYVIEDTWYGYGLLQDSGSLQPEVDYTRPPQGTGVCVPSPT	780
QY	781	SRAPAKLATEEPPRSPSESESEPEYTSHPNNGDVSSKDTDIPILSTQEPME	836
DB	781	SRAPAKLATEEPPRSPSESESEPEYTSHPNNGDVSSKDTDIPILSTQEPME	836

RESULT 2
US-09-899-569A-4
; Sequence 4, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweitzer
; APPLICANT: Marwa Scherl-Mostagier
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher

THIS PAGE IS BLANK

TITLE OF INVENTION: Tumoraassoziiertes Antigen (B345)
FILE REFERENCE: 0652.2280001
CURRENT APPLICATION NUMBER: US/09/899,569A
CURRENT FILING DATE: 2003-07-06
PRIOR APPLICATION NUMBER: DE 100 33 080.0
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: DE 101 19 294.0
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 60/243,158
PRIOR FILING DATE: 2000-10-25
PRIOR APPLICATION NUMBER: US 60/297,747
PRIOR FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 836
TYPE: PRT
ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 99.8%; Score 4385; DB 3; Length 836;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPRGAAFEIALPRESNITVLIKAGPTTLAKPCYIYI 60
DB 1 MAGNCGVSIALGLVLLGAARLPRGAAFEIALPRESNITVLIKAGPTTLAKPCYIYI 60
QY 61 SKRHITMLSTISGERIVTFTSCQSPENHFVIEIQNIDCMGSCPFGGEVOLQSTSLPT 120
DB 61 SKRHITMLSTISGERIVTFTSCQSPENHFVIEIQNIDCMGSCPFGGEVOLQSTSLPT 120
QY 121 LNRFTIMDVAKAKSIGLELOPSIPRLRQIGGESCPDVTSHISGRIDATVYRIGTFCSN 180
DB 121 LNRFTIMDVAKAKSIGLELOPSIPRLRQIGGESCPDVTSHISGRIDATVYRIGTFCSN 180
QY 181 GTVSRIKQEGVKXALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIKQEGVKXALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGFPEDEIMTQVVPVPAHLAASVFLNFNISNCERKEERYEYIIPGSTNPEVFKLEDK 300
DB 241 PEGFPEDEIMTQVVPVPAHLAASVFLNFNISNCERKEERYEYIIPGSTNPEVFKLEDK 300
QY 301 QPGMAGNFNLISLQSCDQDAOSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLISLQSCDQDAOSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKOSRKFPVPGFVCLERTCSNLTLSGSKHKSIFLCDDLTRLMNVEXTISCTDH 420
DB 361 PRPVKOSRKFPVPGFVCLERTCSNLTLSGSKHKSIFLCDDLTRLMNVEXTISCTDH 420
QY 421 RYCORKSYSIQVPSDILHLVEHLDFSKLVPKORLSLVLPKQKLOHTEKPCNTSF 480
DB 421 RYCORKSYSIQVPSDILHLVEHLDFSKLVPKORLSLVLPKQKLOHTEKPCNTSF 480
QY 481 SYLVASAIPEQDLYFGSFCPGSGIKOIQVKNISVTLRTAPSPFORASROGLTVSFIPIY 540
DB 481 SYLVASAIPEQDLYFGSFCPGSGIKOIQVKNISVTLRTAPSPFORASROGLTVSFIPIY 540
QY 541 FKEEGVFTVTPDTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFKERSGVVCO 600
DB 541 FKEEGVFTVTPDTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFKERSGVVCO 600
QY 601 TGRAPMIIQORTARAEIIFSLDEDVLPKPSFHHSFVNI SNCSPTSGKQDLDFSVTLT 660
DB 601 TGRAPMIIQORTARAEIIFSLDEDVLPKPSFHHSFVNI SNCSPTSGKQDLDFSVTLT 660
QY 661 PRIVD.LVILIAVGGVLLLSALGLIICVKKKKKTKNGPAVGIYNGINTEMPROPK 720
DB 661 PRIVD.LVILIAVGGVLLLSALGLIICVKKKKKTKNGPAVGIYNGINTEMPROPK 720
QY 721 KFOGKRKNDSHVYAVIEDTMVYGHLLQSSGSLQPEVDYTRPFGQTMGVCPSPPTIC 780

DB 721 KFOGKRKNDSHVYAVIEDTMVYGHLLQSSGSLQPEVDYTRPFGQTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPPESESEBYTFSHPNNGVSSKQTDIPLISTOEPMEPAP 836
DB 781 SRAPTAKLATEEPPRSPPESESEBYTFSHPNNGVSSKQTDIPLISTOEPMEPAP 836

Search completed: May 4, 2006, 21:05:31
Job time: 168 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 20:53:55 ; Search time 237 Seconds
(without alignments)
2488.699 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4392

Sequence: 1 MAGLNCGVSTALLGVLLIGA.....SSKDTDIPLSTQPMPEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_crembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	2	Q9HSV8_HUMAN
2	4385	99.8	836	2	Q96QU7_HUMAN

ALIGNMENTS

RESULT 1
Q9HSV8_HUMAN PRELIMINARY; PRT; 836 AA.
AC Q9HSV8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein FLJ22969 (NCSG135).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
OX NCBI
RN Nucleotide Sequence.
RA Watanabe K., Kumagai A., Itakura S., Yamashita M., Tashiro H., Ota T.,
RA Suzuki Y., Ohnishi M., Nishii T., Shibahara T., Tanaka T.,
RA Nakamura Y., Iino T., Sugano S.;
RU Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN Nucleotide Sequence.
RP MEDLINE=22547370; PubMed=12660814; DOI=10.1038/sj.onc.1206220;
RA Hooper J.D., Zijlstra A., Aimes R.T., Liang H., Claassen G.F.,
RA Tarin D., Teeta J.E., Quigley J.P.;

RT "Subtractive immunization using highly metastatic human tumor cells
RT identifier SIMA135/CDCL1, a 135 kDa cell surface phosphorylated
RT glycoprotein antigen."
RL Oncogene 22:1783-1794 (2003).
DR EMBL; AK026622; BAB15511.1; -; mRNA.
DR EMBL; AF468010; AA033397.1; -; mRNA.
SQ SEQUENCE 836 AA; 92874 MM; 98980475C35C4C8 CRC64;

Query Match 100.0%; Score 4392; DB 2; Length 836;
Best local similarity 100.0%; Pred. No. 2.4e-312;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAGLNCGVSTALLGVLLIGAARLRPGAEAPIALPRESNTVLKIGPTLLAPRCYVI	60
Db	1	MAGLNCGVSTALLGVLLIGAARLRPGAEAPIALPRESNTVLKIGPTLLAPRCYVI	60
Qy	61	SKRHITMLSIKSGRIYFTSCQSPENHFVIEIOKNDICMSGPCPCFQVQLOPSTSLPT	120
Db	61	SKRHITMLSIKSGRIYFTSCQSPENHFVIEIOKNDICMSGPCPCFQVQLOPSTSLPT	120
Qy	121	LNRFTIWDVKAHKSIGLEQPSIPRLRQIGPSSCPDGVTHSIGRIDATVVRIGTFCSN	180
Db	121	LNRFTIWDVKAHKSIGLEQPSIPRLRQIGPSSCPDGVTHSIGRIDATVVRIGTFCSN	180
Qy	181	GTYSRIKQSGVKKALHPWHPNNVSGFSIANSSIKRLCIISVREGSSATLMSANY	240
Db	181	GTYSRIKQSGVKKALHPWHPNNVSGFSIANSSIKRLCIISVREGSSATLMSANY	240
Qy	241	PEGFPEDELMTWQVFPVAHLRASVPLNPNLSNCRKEERYEYIIPSTTPPEVFKLEDK	300
Db	241	PEGFPEDELMTWQVFPVAHLRASVPLNPNLSNCRKEERYEYIIPSTTPPEVFKLEDK	300
Qy	301	QPGMAGNFNLSLQCGDQDQSPGILRLQFVLVQHPQNSNKIYVVDLSNERAMSLTIE	360
Db	301	QPGMAGNFNLSLQCGDQDQSPGILRLQFVLVQHPQNSNKIYVVDLSNERAMSLTIE	360
Qy	361	PRPVKQSRKRVPGFVCLBSRTGSSNLTLSGSHKISFLCDDLTRLMANVEKTIISCTDH	420
Db	361	PRPVKQSRKRVPGFVCLBSRTGSSNLTLSGSHKISFLCDDLTRLMANVEKTIISCTDH	420
Qy	421	RYCQRKSYSLQVPSDILHLPELHDFSMKLLVPKDRSLVLPVPAQKIQDHTHEPCNTSF	480
Db	421	RYCQRKSYSLQVPSDILHLPELHDFSMKLLVPKDRSLVLPVPAQKIQDHTHEPCNTSF	480
Qy	481	SYLVASAIIPQDLYFGSFCGSGIKQIQVQKNSIVTLRTFAPSROBASRQGLTVSEFIPY	540
Db	481	SYLVASAIIPQDLYFGSFCGSGIKQIQVQKNSIVTLRTFAPSROBASRQGLTVSEFIPY	540
Qy	541	FKERGVTVPDRTSKYVLRTPWDRGLPSLTYSNMI SVPRDVAALTFPKKSGVYVQ	600
Db	541	FKERGVTVPDRTSKYVLRTPWDRGLPSLTYSNMI SVPRDVAALTFPKKSGVYVQ	600
Qy	601	TGRAFMIIORPRAEITFSLDEVDLPKPSFHHHSFWNVSNCSPYSGKQDLDFSTLT	660
Db	601	TGRAFMIIORPRAEITFSLDEVDLPKPSFHHHSFWNVSNCSPYSGKQDLDFSTLT	660
Qy	661	PRIVDLTVIIIAVGGVLLLSALGLIICVKKKKKTKNGPAVINGNINTEMPROPK	720
Db	661	PRIVDLTVIIIAVGGVLLLSALGLIICVKKKKKTKNGPAVINGNINTEMPROPK	720
Qy	721	KFOGRKNDNSHYAVVEDTMVYGHLLQDSGSGFLQPEVDYTRPFQGTMGVCPSPPTIC	780
Db	721	KFOGRKNDNSHYAVVEDTMVYGHLLQDSGSGFLQPEVDYTRPFQGTMGVCPSPPTIC	780
Qy	781	SRAPTALATEBPPEPSPSESEPTFFSHPNNDVSKOTDIPLSTQPMPEPAE 836	
Db	781	SRAPTALATEBPPEPSPSESEPTFFSHPNNDVSKOTDIPLSTQPMPEPAE 836	

RESULT 2
Q96QU7_HUMAN PRELIMINARY; PRT; 836 AA.
AC Q96QU7;

DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CUB domain containing protein 1.
GN Name=CDCL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21359860; PubMed=11466221; DOI=10.1038/sj.onc.1204566;
RA Scherl-Mostlager M., Sommergruber W., Abseher R., Hauptmann R.,
RA Ambros P., Schewelfer N.,
RT "Identification of a novel gene, CDCL1, overexpressed in human
RT colorectal cancer."
RL Oncogene 20:4402-4408(2001).
DR EMBL; AY026461; AAK02058.1; -; mRNA.
DR Ensembl; ENSG00000163814; Homo sapiens.
SQ SEQUENCE 836 AA; 92873 MW; FB4D2DBBD35C519 CRC64;

Query Match 99.8%; Score 4385; DB 2; Length 836;
Best Local Similarity 99.8%; Pred. No. 7.8e-312;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCVSILLGLGALPGARAFETALPRESNITVLKLGPTLLAKPCYVI 60
DB 1 MAGINCVSILLGLGALPGARAFETALPRESNITVLKLGPTLLAKPCYVI 60
QY 61 SKRHITMLSIKGERIVFTFSCSPENHFVIEIQKINIDCMGSPCPFGEVOLQPSLPT 120
DB 61 SKRHITMLSIKGERIVFTFSCSPENHFVIEIQKINIDCMGSPCPFGEVOLQPSLPT 120
QY 121 LNRFTIWDVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKMALHPMFHPRNVSGFSIANRSSIKRLCTIESVPEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVKMALHPMFHPRNVSGFSIANRSSIKRLCTIESVPEGEGSATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRSVFLNPNLSNCRKEERVEYIIPGSTNPEVFKLEBK 300
DB 241 PEGFPEDELMTQFVVPALHRSVFLNPNLSNCRKEERVEYIIPGSTNPEVFKLEBK 300
QY 301 QPGNAGNPNLSLQGCDDAOSPGILRLQFOYLVQHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNAGNPNLSLQGCDDAOSPGILRLQFOYLVQHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSKKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTSCTDH 420
DB 361 PRPVQSKKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTSCTDH 420
QY 421 RYCQKSKSYLQVPSDILHLPELHDFSWKLVPKDRLSLVLPKQLQOHTHEKPCNTSF 480
DB 421 RYCQKSKSYLQVPSDILHLPELHDFSWKLVPKDRLSLVLPKQLQOHTHEKPCNTSF 480
QY 481 SYLVASAIIPSDLLYGSFCPGSGIKOIQVKONISVTLRTFAPSFOEASROGLTVSFIPIY 540
DB 481 SYLVASAIIPSDLLYGSFCPGSGIKOIQVKONISVTLRTFAPSFOEASROGLTVSFIPIY 540
QY 541 FKEBGVFTVPTPKSKVLRTPNMDRGLPSLTSVSMNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGVFTVPTPKSKVLRTPNMDRGLPSLTSVSMNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAFMITQORTRAEEIFSLDEDEVLPKPSFHHHSFWVNISNCSPTSQKQDLDFSVTLT 660
DB 601 TGRAFMITQORTRAEEIFSLDEDEVLPKPSFHHHSFWVNISNCSPTSQKQDLDFSVTLT 660
QY 661 PRTVDLTVILLAAVGGVLLLSALGLIICVKKKKKTKNGPAVGIIYNGNINTEMPROPK 720
DB 661 PRTVDLTVILLAAVGGVLLLSALGLIICVKKKKKTKNGPAVGIIYNGNINTEMPROPK 720

QY 721 KEQGRKNDSSHVYAVIEDITWVYGHLLQDSSGSFLQPEVDTRPPFOGTMGVCPSPPTIC 780
DB 721 KEQGRKNDSSHVYAVIEDITWVYGHLLQDSSGSFLQPEVDTRPPFOGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPPESESSEPPYTFSHPNNGDVSSKDTDIPLLSTQEBMEPAE 836
DB 781 SRAPTAKLATEEPPRSPPESESSEPPYTFSHPNNGDVSSKDTDIPLLSTQEBMEPAE 836

Search completed: May 4, 2006, 21:00:51
Job time : 238 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:05:42 ; Search time 134 Seconds
(without alignments)
2741.199 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4377

Sequence: 1 MAGUNCGSIALGLVLLGA.....SSKDTDPLLTQEMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_21:*
- 1: geneseqp1980s:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*
 - 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4375	99.9	836	ADR47303	Adt47303 Human SIM
2	4375	99.9	836	Aea01805	Aea01805 Human CUB
3	4374	99.9	836	ABP69553	ABP69553 Human POL
4	4374	99.9	836	AAM49641	AAM49641 Human tum
5	3625.5	82.8	709	AAV91456	AAV91456 Human sec
6	3625.5	82.8	709	ADL71528	ADL71528 Human hum
7	3454	78.9	749	AAM49640	AAM49640 Human tum
8	3395	77.6	649	AAV95558	AAV95558 Human pro
9	2155	49.2	414	ADL71669	ADL71669 Novel hum
10	2155	49.2	414	AAV91592	AAV91592 Human sec
11	2155	49.2	443	ADL71667	ADL71667 Novel hum
12	1779	40.6	343	AAU12252	AAU12252 Human PRO
13	1779	40.6	343	ABD17686	ABD17686 Novel hum
14	1779	40.6	343	ABU80950	ABU80950 Human PRO
15	1779	40.6	343	ABU66650	ABU66650 Human PRO
16	1779	40.6	343	ABU59731	ABU59731 Novel sec
17	1779	40.6	343	ABO24921	ABO24921 Human sec
18	1779	40.6	343	ABU66926	ABU66926 Human sec
19	1779	40.6	343	ADA56581	ADA56581 Novel hum
20	1779	40.6	343	ADA76112	ADA76112 Human PRO
21	1779	40.6	343	ADA18762	ADA18762 Human PRO
22	1779	40.6	343	ADA61385	ADA61385 Homo sapi
23	1779	40.6	343	ADB19170	ADB19170 Novel hum
24	1779	40.6	343	ADB27711	ADB27711 Human PRO

25	1779	40.6	343	6	ADA86190	ADA86190 Novel hum
26	1779	40.6	343	6	ADB15754	ADB15754 Human PRO
27	1779	40.6	343	6	ADA47540	ADA47540 Human PRO
28	1779	40.6	343	6	ADA67335	ADA67335 Human PRO
29	1779	40.6	343	6	ADB30342	ADB30342 Human PRO
30	1779	40.6	343	6	ADA85638	ADA85638 Novel hum
31	1779	40.6	343	6	ADA96850	ADA96850 Human PRO
32	1779	40.6	343	6	ADA79154	ADA79154 Human PRO
33	1779	40.6	343	6	ADA87293	ADA87293 Novel hum
34	1779	40.6	343	6	ADB16495	ADB16495 Human PRO
35	1779	40.6	343	6	ADA91587	ADA91587 Novel hum
36	1779	40.6	343	6	ADB14650	ADB14650 Human PRO
37	1779	40.6	343	6	ADB19611	ADB19611 Novel hum
38	1779	40.6	343	6	ADA93826	ADA93826 Human PRO
39	1779	40.6	343	6	ADB19722	ADB19722 Novel hum
40	1779	40.6	343	6	ADB13034	ADB13034 Human PRO
41	1779	40.6	343	6	ABO43229	ABO43229 Novel hum
42	1779	40.6	343	6	ADA74288	ADA74288 Human PRO
43	1779	40.6	343	6	ADB24521	ADB24521 Human PRO
44	1779	40.6	343	6	ADA82045	ADA82045 Human PRO
45	1779	40.6	343	6	ADA75008	ADA75008 Human PRO

ALIGNMENTS

RESULT 1
ADR47303
ID ADR47303 standard; protein; 836 AA.
XX
AC ADR47303;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human SIMA135 protein.
XX
KW tumour marker protein; glycosylated; non-glycosylated; antibody;
KM cytosolic; cancer; metastasis; SIMA135; human.
XX
OS Homo sapiens.
XX
PN WO2004074481-A1.
XX
PD 02-SEP-2004.
XX
PF 18-FEB-2004; 2004WO-EP001556.
XX
PR 19-FEB-2003; 2003US-0448828P.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS PHARMA GMBH.
PA (SCRI) SCRIPPS RES INST.
XX
PI Quigley JP, Hooper JD, Testa JE.
XX
WP1: 2004-652958/63.
XX
PT New glycoprotein antigen SIMA135, useful for diagnosing and treating
PT cancer, e.g. prostate, colon, gastric, liver, breast, lung, or kidney
PT cancer.
XX
PS Claim 1: SEQ ID NO 1; S1pp; English.
XX
The invention relates to a novel tumour marker protein. The protein
comprises a sequence of 836 amino acids, ADR47303, or its fragment or
variant, that is glycosylated or non-glycosylated, where the variant has
an amino acid at position 525 which is Arginine and/or the amino acid at
position 827 is Serine. The invention further comprises: an antibody that
binds specifically to the protein, or its fragment or variant, where the
antibody is not mAb 41-2; a pharmaceutical composition comprising the
antibody and a pharmaceutical carrier; and a kit comprising the antibody
and packaging material. The novel protein, its fragments and variants,
and the antibody have cytostatic activity. The antibody is used in a

CC diagnostic method of cancer, to determine if a test sample contains
CC metastatic cells, a therapeutic treatment of the human and animal body,
CC to determine the metastasis modulating ability of an agent, and to
CC determine if a candidate agent modulates SIMA15 production by a cell.
CC The antibody is useful for preparing a medicament to inhibit metastasis
CC by a cancer cell, e.g. epidermoid carcinoma cell, a fibrosarcoma, or a
CC prostate, colon, gastric, liver, breast, lung, or kidney rhabdoid cancer
CC cell, or Hep3 cell, in a mammal. The protein or its fragment or variant,
CC method, and composition are also useful in diagnosing and treating
CC cancer. This sequence represents a human SIMA15 protein of the
CC invention.

CC
XX
SQ Sequence 836 AA;

Query Match 99.9%; Score 4375; DB 8; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPFGAFAEIALPRESNITVLKLGPTLLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPFGAFAEIALPRESNITVLKLGPTLLAKPCYIYI 60
QY 61 SKRHITMLSIKSGERIVFTFSCSPENHFVIEIOKNIDMSGPCPEGEVQLQSTSLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCSPENHFVIEIOKNIDMSGPCPEGEVQLQSTSLPT 120
QY 121 LNRITIMVYKAHKSIGLEIFSIPLRQIGPESCPDGYTHISGRIDATVYRIGTFCSN 180
DB 121 LNRITIMVYKAHKSIGLEIFSIPLRQIGPESCPDGYTHISGRIDATVYRIGTFCSN 180
QY 181 GTVSRITKQEGVMALHLPMFAPRVNVSFSIANRSIKLCITIEFEGEGSATLMSANY 240
DB 181 GTVSRITKQEGVMALHLPMFAPRVNVSFSIANRSIKLCITIEFEGEGSATLMSANY 240
QY 241 PEGPEDEIMTQFVVPALIRASVSFLNENLSCNCKEKEVEYIIPGSTTNEVFLEDK 300
DB 241 PEGPEDEIMTQFVVPALIRASVSFLNENLSCNCKEKEVEYIIPGSTTNEVFLEDK 300
QY 301 QPQNMAGNPNLSLOGCDODAGSPGILRLOFQVLYQHPOESNKIYVVDLSNRAMSLTE 360
DB 301 QPQNMAGNPNLSLOGCDODAGSPGILRLOFQVLYQHPOESNKIYVVDLSNRAMSLTE 360
QY 361 PAPVQSRKFVPGCFVCSRTCSNLTLSGSKHKISFLCDLFTLMANVEKITSCTDH 420
DB 361 PAPVQSRKFVPGCFVCSRTCSNLTLSGSKHKISFLCDLFTLMANVEKITSCTDH 420
QY 421 RYCQKRSYSLOVPSDILHLPELHDFSMLVPCORLSLVLVPAQLOQHTHEKPCNTSF 480
DB 421 RYCQKRSYSLOVPSDILHLPELHDFSMLVPCORLSLVLVPAQLOQHTHEKPCNTSF 480
QY 481 SYLVASATPSODLYRSGPCPGSIKOIYQKONISVTLRFPASFOEASROGLTVSFIY 540
DB 481 SYLVASATPSODLYRSGPCPGSIKOIYQKONISVTLRFPASFOEASROGLTVSFIY 540
QY 541 FPEBEGFTVTPDKSKVYLRTPNMDGLPSLTSVSNISVPRDQVACLTFFERSGVQO 600
DB 541 FPEBEGFTVTPDKSKVYLRTPNMDGLPSLTSVSNISVPRDQVACLTFFERSGVQO 600
QY 601 TGRAPMTIOEORTRAEIFSLDEDEVLPKPSFHHSFWANISNCSPTSGKQDLLESVLT 660
DB 601 TGRAPMTIOEORTRAEIFSLDEDEVLPKPSFHHSFWANISNCSPTSGKQDLLESVLT 660
QY 661 PRTVULTVLLIANGGVLLSALGLITICVKKKKKKTKKGAVGVYNNINTEMPROK 720
DB 661 PRTVULTVLLIANGGVLLSALGLITICVKKKKKKTKKGAVGVYNNINTEMPROK 720
QY 721 KFOKGRKNDSHVYAVIEDTMVYGHLLQDSGSGFLOPEVDYTRPFGTGVCPSPPTIC 780
DB 721 KFOKGRKNDSHVYAVIEDTMVYGHLLQDSGSGFLOPEVDYTRPFGTGVCPSPPTIC 780
QY 781 SPAPTAKLATEBPPRSPSESESEPTFSHPNNGDVSSKOTDIPILXTOEPMEPAR 836
DB 781 SPAPTAKLATEBPPRSPSESESEPTFSHPNNGDVSSKOTDIPILXTOEPMEPAR 836

RESULT 2
AEA01805
ID AEA01805 standard; protein, 836 AA.

XX AEA01805;
AC 14-UTL-2005 (first entry)
XX

DB Human CUB domain containing protein 1 (CDCP1).

XX CUB domain containing protein 1; CDCP1; ovary tumor; cytostatic;
KW endocrine disease; genitourinary disease; gynecology and obstetrics;
KW neoplasm; gene therapy; vaccine; diagnosis; drug screening.

OS Homo sapiens.

FT Key Location/Qualifiers
FT Domain 30..667
FT /note="Predicted extracellular domain"

PN WO2005042102-A1.

PD 12-MAY-2005.

PE 22-OCT-2004; 2004WO-GB004502.

PR 22-OCT-2003; 2003GB-00024656.

XX (CLLT) CELLTech R & D LTD.

PI Burgess NA;

XX WPI, 2005-366521/37.

DR N-PSDB; AEA01806.

PT Use of CDCP1 polypeptide or an agent that interacts with or modulates the
PT expression or activity of CDCP1 polypeptide for diagnosing, preventing or
PT treating ovarian cancer or in screening for agents that may treat or
PT prevent the disease.

PS Claim 6; SEQ ID NO 1; 48bp; English.

XX The present sequence is the protein sequence of human CUB domain
CC containing protein 1 (CDCP1). The invention is based on the finding that
CC CDCP1 represents a novel therapeutic target for the treatment and/or
CC prophylaxis of ovarian cancer. The invention relates to new uses of CDCP1
CC in the diagnosis, screening, treatment and prophylaxis of ovarian cancer,
CC and provides compositions comprising CDCP1, antibodies that are
CC immunospecific for CDCP1, and agents that interact with or modulate the
CC expression or activity of CDCP1 or which modulate the expression a
CC nucleic acid encoding CDCP1. The CDCP1 polypeptide comprises the present
CC sequence or is a derivative of this sequence. The use of an agent that
CC interacts with or modulates the expression or activity of a CDCP1
CC polypeptide for the manufacture of a medicament for the treatment and/or
CC prophylaxis of ovarian cancer is claimed. The agent may be an antibody or
CC its functionally active fragment, derivative or analog, which may be
CC conjugated to a therapeutic moiety, detectable label, second antibody or
CC fragment, effector or reporter molecule, cytotoxic agent or cytokine. A
CC claimed method for the treatment and/or prophylaxis of ovarian cancer
CC involves administering a composition comprising a CDCP1 polypeptide. The
CC composition may be a vaccine. Also claimed are methods of using CDCP1 to
CC screen for anti-ovarian cancer agents, and a method of screening for
CC and/or diagnosis or prognosis of ovarian cancer in a subject, and/or
CC monitoring the effectiveness of ovarian cancer therapy in a subject by
CC detecting and/or quantifying expression of CDCP1 in a sample obtained
CC from the subject.

SO Sequence 836 AA;

Query Match 99.9%; Score 4375; DB 9; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;

Matches	833: Conservative	0: Mismatches	3: Indels	0: Gaps	0:
QY	1	MAGNCGVSIALLGVLLGGAARLPRGAFAEIALPRESNTITVLKLGPTLLAPCYVI	60		
Db	1	MAGNCGVSIALLGVLLGGAARLPRGAFAEIALPRESNTITVLKLGPTLLAPCYVI	60		
QY	61	SKRHITMISIKSGRIIVFTFSCQSPENHFVEIQKNIDCMGSPCFGEVQLQPSLSLPT	120		
Db	61	SKRHITMISIKSGRIIVFTFSCQSPENHFVEIQKNIDCMGSPCFGEVQLQPSLSLPT	120		
QY	121	LNRTFIWVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180		
Db	121	LNRTFIWVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180		
QY	181	GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVGECSATLMSANY	240		
Db	181	GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVGECSATLMSANY	240		
QY	241	PEGFPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPEVFKLBDK	300		
Db	241	PEGFPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPEVFKLBDK	300		
QY	301	QPGNAGNPNLSLQSCDDAOSPGLRLQFVLVQHPONESNKIYVDLSNERAMSLTIE	360		
Db	301	QPGNAGNPNLSLQSCDDAOSPGLRLQFVLVQHPONESNKIYVDLSNERAMSLTIE	360		
QY	361	PRPVQSRKFPVPGCFVCLBSRTCSNLTITSGSKHKISFLCDLTRLNMANVEKTIISCTDH	420		
Db	361	PRPVQSRKFPVPGCFVCLBSRTCSNLTITSGSKHKISFLCDLTRLNMANVEKTIISCTDH	420		
QY	421	RYCORKSYSLQVPSDILHLPEVLHDFSMKLVLPKRLSLVLPACKLOOHTHEKCNISF	480		
Db	421	RYCORKSYSLQVPSDILHLPEVLHDFSMKLVLPKRLSLVLPACKLOOHTHEKCNISF	480		
QY	481	SYLVASAIPODLVGFSCPGSGIKQIVKONISVTLRTFAPSEFYQASRQGLTVPFIPY	540		
Db	481	SYLVASAIPODLVGFSCPGSGIKQIVKONISVTLRTFAPSEFYQASRQGLTVPFIPY	540		
QY	541	FKESEVFPVTPTTSKVTLRTPNMDRGLPSLTVSWMNISVPRDOVACLTFEKERSGVVQC	600		
Db	541	FKESEVFPVTPTTSKVTLRTPNMDRGLPSLTVSWMNISVPRDOVACLTFEKERSGVVQC	600		
QY	601	TGRAFMIIQEOBRTAREEIFSDEVDLPKPSFHHHSFWNINSCSTSCOKDILFSTYLT	660		
Db	601	TGRAFMIIQEOBRTAREEIFSDEVDLPKPSFHHHSFWNINSCSTSCOKDILFSTYLT	660		
QY	661	PRVDLTIVILIAAVGGVLLSALGLIICVYKKKKKKTKNKGPAGVITVNGNINTEMPROPK	720		
Db	661	PRVDLTIVILIAAVGGVLLSALGLIICVYKKKKKKTKNKGPAGVITVNGNINTEMPROPK	720		
QY	721	KFOGRKNDKSHVAVIEDTWYGHLLQDSSGSLQPEVDVTRPRFQGTMGVCPSPPTIC	780		
Db	721	KFOGRKNDKSHVAVIEDTWYGHLLQDSSGSLQPEVDVTRPRFQGTMGVCPSPPTIC	780		
QY	781	SRAPIPAKATBEPPRPSPSESESEPYTSHPNNGVSSKDPIDILPLSTQEPWEPAP	836		
Db	781	SRAPIPAKATBEPPRPSPSESESEPYTSHPNNGVSSKDPIDILPLSTQEPWEPAP	836		

RESULT 3
ABP69553
ID ABP69553 standard; protein, 836 AA.
XX ABP69553;
DT 20-JAN-2003 (first entry)
XX Human polypeptide SEQ ID NO 1600.
DE Human, genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;

KW	arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW	haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
XX	antiarthritic.
OS	Homo sapiens.
XX	WO2002/0539-A2.
PN	12-SEP-2002.
PD	05-MAR-2002; 2002WO-US005095.
PF	05-MAR-2002; 2002WO-US005095.
XX	05-MAR-2001; 2001US-00799451.
PR	(HYSE-) HYSEQ INC.
XX	Tang YF, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI	Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI	Wehrman T, Wang J, Wang D, Dmanac RT;
XX	WPI; 2002-759812/82.
DR	N-PSDB; AB211770.
XX	New polynucleotides comprising sequences assembled from expressed
PT	sequence tags (ESTs), useful for treating cell-proliferative,
PT	neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT	or coagulation disorders.
XX	Claim 9; SEQ ID NO 1600; 1012pp + Sequence Listing; English.
PS	The invention relates to an isolated polynucleotide (I) comprising a
CC	nucleotide sequence selected from any of 948 sequences (AB21119-
CC	AB212066) or their mature protein coding portion, active domain coding
CC	protein or complementary sequences. The polynucleotides are useful for
CC	identifying expressed genes or for physical mapping of human genome. The
CC	encoded polypeptides (ABP6902-ABP69849) are useful as molecular weight
CC	markers, as a food supplement, for generating antibodies, in medical
CC	imaging, screening and diagnostic assays and for treating cell-
CC	proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC	or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC	dibetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC	platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC	or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC	arthritis, etc. Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic format
CC	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	Sequence 836 AA:
QY	1 MAGNCGVSIALLGVLLGGAARLPRGAFAEIALPRESNTITVLKLGPTLLAPCYVI
Db	1 MAGNCGVSIALLGVLLGGAARLPRGAFAEIALPRESNTITVLKLGPTLLAPCYVI
QY	61 SKRHITMISIKSGRIIVFTFSCQSPENHFVEIQKNIDCMGSPCFGEVQLQPSLSLPT
Db	61 SKRHITMISIKSGRIIVFTFSCQSPENHFVEIQKNIDCMGSPCFGEVQLQPSLSLPT
QY	121 LNRTFIWVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN
Db	121 LNRTFIWVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN
QY	181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVGECSATLMSANY
Db	181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVGECSATLMSANY
QY	241 PEGFPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPEVFKLBDK
Db	241 PEGFPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPEVFKLBDK

Query Match 99.9%; Score 4374; DB 5; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 301 OPGNMGAFNLSLQCCDDAOSPGILRLQFVLYVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 OPGNMGAFNLSLQCCDDAOSPGILRLQFVLYVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSKRFVPGCFVCLIESRTCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
DB 361 PRPVQSKRFVPGCFVCLIESRTCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
QY 421 RYCQKKSYSLOVPSDILHLPEVLHDFSWKLLVPPKORLSLVVPAQKLOQTHEKPCNTSF 480
DB 421 RYCQKKSYSLOVPSDILHLPEVLHDFSWKLLVPPKORLSLVVPAQKLOQTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYFGSFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIPY 540
DB 481 SYLVASAIPSODLYFGSFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIPY 540
QY 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTYSVSNISVPRDQVACLTFFKERSGVVQ 600
DB 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTYSVSNISVPRDQVACLTFFKERSGVVQ 600
QY 601 TGRAFMIIOBQRTABEIFSLEDEVLPKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
DB 601 TGRAFMIIOBQRTABEIFSLEDEVLPKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
QY 661 PRTVDLTVILIAAVGGVLLLSALGLIICVYKKKKKTKNGPANGIYNNXNITEMPROPK 720
DB 661 PRTVDLTVILIAAVGGVLLLSALGLIICVYKKKKKTKNGPANGIYNNXNITEMPROPK 720
QY 721 KFQGRKONDSHVYAVIEDTMYGHLLODSGSGFLQPEVDYTRPFGQTMGVCPPSPPTIC 780
DB 721 KFQGRKONDSHVYAVIEDTMYGHLLODSGSGFLQPEVDYTRPFGQTMGVCPPSPPTIC 780
QY 781 SRAPTAKLATEBPSPRSPESESEPTFSHPNNGDVSSKOTDIPLIXTOEPMEPAE 836
DB 781 SRAPTAKLATEBPSPRSPESESEPTFSHPNNGDVSSKOTDIPLIXTOEPMEPAE 836

RESULT 4
AAM49641
ID AAM49641 standard; protein, 836 AA.
XX
AC AAM49641;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human tumour-associated antigen B345 protein SEQ ID NO 4.
XX
KW Tumour-associated antigen; human; B345; cytostatic; cell communication;
KW cell interaction; signal transduction; metastasis; cancer; colon;
KW immunotherapy; carcinoma; lung; diagnosis.
XX
OS Homo sapiens.
XX
PN WO200204508-A1.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-BP007705.
XX
PR 07-JUL-2000; 2000DE-01033080.
XX
PR 19-APR-2001; 2001DE-01019294.
XX
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI Schweiler N, Scherl-Mostageer M, Sommergruber W, Abseher R;
XX
DR WPI: 2002-171704/22.
XX
DR N-PSDB; ABA95507.
XX
PT New tumor-associated antigen B345, useful for diagnosis and immunotherapy
PT of tumors, also related nucleic acid and antibodies.
XX
```

```
PS Claim 1; Page 85-88; 102pp; German.
XX
CC This invention describes a novel tumour-associated antigen, designated
CC B345 which has cytostatic activity. B345 is involved in communication,
CC interaction and/or signal transduction with extracellular components and
CC ligands, especially in the metastatic potential of cancers, particularly
CC of the colon. B345 or its immunogenic fragments, also the DNA that
CC encodes it, are useful for immunotherapy of cancer, particularly
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for
CC treatment and diagnosis of cancers that are associated with B345
CC expression, including their use for targeted delivery of cytotoxic or
CC radioactive agents. Probes derived from B345 can be used to detect tumour
CC -specific mutations in the B345 sequence, and can be used to screen for
CC B345 specific modulators. This sequence represents a human B345 tumour-
CC associated antigen described in the invention
SQ
XX
Sequence 836 AA;
Query Match 99.9%; Score 4374; DB 5; Length 836;
Beet Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAGLNCGVSIALLGVLLGAARLPAGAEAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAARLPAGAEAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
QY 61 SKRHITMISTSGEIVTFSCQSENFHVEIQNDICMGSPCFGEVQLOPSTSLPT 120
DB 61 SKRHITMISTSGEIVTFSCQSENFHVEIQNDICMGSPCFGEVQLOPSTSLPT 120
QY 121 INRFTIMDKAHKSIGLEQFSIPRLQIGGESCPEQVTHSISGRIDATVIRIGTFCSN 180
DB 121 INRFTIMDKAHKSIGLEQFSIPRLQIGGESCPEQVTHSISGRIDATVIRIGTFCSN 180
QY 181 GTVSRIKMOEGVKALHLPMFHPRNVSQFSIANNSSIRLCIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVKALHLPMFHPRNVSQFSIANNSSIRLCIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELTMOQFVVPAPHLRASVSFLNPNINSNCERKEERYVYRGSTTNPEVKLEBK 300
DB 241 PEGFPEDELTMOQFVVPAPHLRASVSFLNPNINSNCERKEERYVYRGSTTNPEVKLEBK 300
QY 301 OPGNMGAFNLSLQCCDDAOSPGILRLQFVLYVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 OPGNMGAFNLSLQCCDDAOSPGILRLQFVLYVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSKRFVPGCFVCLIESRTCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
DB 361 PRPVQSKRFVPGCFVCLIESRTCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
QY 421 RYCQKKSYSLOVPSDILHLPEVLHDFSWKLLVPPKORLSLVVPAQKLOQTHEKPCNTSF 480
DB 421 RYCQKKSYSLOVPSDILHLPEVLHDFSWKLLVPPKORLSLVVPAQKLOQTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYFGSFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIPY 540
DB 481 SYLVASAIPSODLYFGSFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIPY 540
QY 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTYSVSNISVPRDQVACLTFFKERSGVVQ 600
DB 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTYSVSNISVPRDQVACLTFFKERSGVVQ 600
QY 601 TGRAFMIIOBQRTABEIFSLEDEVLPKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
DB 601 TGRAFMIIOBQRTABEIFSLEDEVLPKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
QY 661 PRTVDLTVILIAAVGGVLLLSALGLIICVYKKKKKTKNGPANGIYNNXNITEMPROPK 720
DB 661 PRTVDLTVILIAAVGGVLLLSALGLIICVYKKKKKTKNGPANGIYNNXNITEMPROPK 720
QY 721 KFQGRKONDSHVYAVIEDTMYGHLLODSGSGFLQPEVDYTRPFGQTMGVCPPSPPTIC 780
DB 721 KFQGRKONDSHVYAVIEDTMYGHLLODSGSGFLQPEVDYTRPFGQTMGVCPPSPPTIC 780
```

QY 761 SRAPTAKLATEPPPPSPSESEBPTTSHPNNGVSSKDTPIPLIXTOEPMEPAE 836
 DB 761 SRAPTAKLATEPPPPSPSESEBPTTSHPNNGVSSKDTPIPLIXTOEPMEPAE 836

RESULT 5
 AAY91456
 ID AAY91456 standard; protein: 709 AA.
 XX AAY91456;
 AC AAY91456;
 DT 29-JUN-2000 (first entry)
 XX

Human secreted protein sequence encoded by gene 6 SEQ ID NO:129.

DE Human secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW anti-HIV; antiinflammatory; nocrotropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiashma;
 KW antipneumatic; cardiac; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour.

XX Homo sapiens.
 OS
 PN WO200006698-A1.
 XX
 PD 10-FEB-2000.
 PF 29-JUL-1999; 99WO-US017130.
 XX
 PR 30-JUL-1998; 98US-0094657P.
 PR 05-AUG-1998; 98US-0095486P.
 PR 06-AUG-1998; 98US-0095454P.
 PR 06-AUG-1998; 98US-0095455P.
 PR 12-AUG-1998; 98US-0096319P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 P1 Komatouls GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 P1 Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 P1 Soppet DR, Endress GA, Edner R, Olsen HS, Mucenki M;
 XX
 DR MPI: 2000-195282/17.
 DR N-PSDB: AAA2651.
 XX
 PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancer, neurological
 PT disorders, immune diseases, inflammation or blood disorders.
 XX
 PS Claim 11; Page 456-459; 634pp; English.

XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;
 CC immunosuppressive; anti-HIV; antiinflammatory; nocrotropic; neuroprotective;
 CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
 CC antiashma; antipneumatic; and cardiac. The polynucleotides and their
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
 CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
 CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC reproductive disorders, gastrointestinal disorders, respiratory disorders
 CC and metabolic disorders. The proteins or polynucleotides can also be used

CC as food additives or preservatives. The proteins are also useful for
 CC identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are
 CC sequences used in the exemplification of the present invention
 CC

SQ Sequence 709 AA;
 Query March 82.8%; Score 3625.5; DB 3; Length 709;
 Best Local Similarity 98.3%; Pred. No. 1.9e-308;
 Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAGINCGVSIALLGVLLGAAALPRGAAPFIALPRESNITVTLKGTPTLLAKPCYVI 60
 DB 1 MAGINCGVSIALLGVLLGAAALPRGAAPFIALPRESNITVTLKGTPTLLAKPCYVI 60

QY 61 SKRHITMLISIGSERIVFTFSCQSPENHFVIEIOKNIDCMGCPGEGVOLOPSTSLPT 120
 DB 61 SKRHITMLISIGSERIVFTFSCQSPENHFVIEIOKNIDCMGCPGEGVOLOPSTSLPT 120

QY 121 LNRFTFIMVVKAKHSGIGLEQFSIRPLROIGGECPCDVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTFIMVVKAKHSGIGLEQFSIRPLROIGGECPCDVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKMGEGVKALHLPFHPNRVSGFSIANRSSIKRLCTISVFEGEGSATLMSANY 240
 DB 181 GTVSRIRKMGEGVKALHLPFHPNRVSGFSIANRSSIKRLCTISVFEGEGSATLMSANY 240

QY 241 PEGFPEDELMTWQFVVPALHRAVSFLNFINLNCCKRKEBEVYIIPGSTNPEVFKLEDK 300
 DB 241 PEGFPEDELMTWQFVVPALHRAVSFLNFINLNCCKRKEBEVYIIPGSTNPEVFKLEDK 300

QY 301 QPGMAGNFINSLQCCODDQASPGILRLQPOVVOHPONENKLYVVDLSNERAMSLTIE 360
 DB 301 QPGMAGNFINSLQCCODDQASPGILRLQPOVVOHPONENKLYVVDLSNERAMSLTIE 360

QY 361 PRPVKQSRKFPVPGFCVLESRTCSNLTITSGSKHKSIFLCDDTLRLMNVKXTISCTDH 420
 DB 361 PRPVKQSRKFPVPGFCVLESRTCSNLTITSGSKHKSIFLCDDTLRLMNVKXTISCTDH 420

QY 421 RYCQRKSYSLQVPSDILHLPELHDFSWKLLVPKRLSLVLPAQKLOQHTHEKPCNTSF 480
 DB 421 RYCQRKSYSLQVPSDILHLPELHDFSWKLLVPKRLSLVLPAQKLOQHTHEKPCNTSF 480

QY 481 SYLVASAPSDQLFPGSPCGSGIKQIOVKONISVTLRTFAPSROESRQGLVTSFIPY 540
 DB 481 SYLVASAPSDQLFPGSPCGSGIKQIOVKONISVTLRTFAPSROESRQGLVTSFIPY 540

QY 541 FKESGVFTVPDTSKYVLRTPNMDRGPLSTYSWMNISVPRDOVACLTFFKESGVVCO 600
 DB 541 FKESGVFTVPDTSKYVLRTPNMDRGPLSTYSWMNISVPRDOVACLTFFKESGVVCO 600

QY 601 TGRAFMIIQEQRTAEIIFSLDEEDVLPKPSFHHHSFWVNISNCSPTSGKQDLPLFSVTLT 660
 DB 601 TGRAFMIIQEQRTAEIIFSLDEEDVLPKPSFHHHSFWVNISNCSPTSGKQDLPLFSVTLT 660

QY 661 PRYDVLVYILAAVGGVLLSALGLITICVKKKKKTKNKPAVGIYXN 710
 DB 661 PRYDVLVYILAAVGGVLLSALGLITICVKKKKKTKNKPAVGIYXN 710

RESULT 6
 ADL71528
 ID ADL71528 standard; protein: 709 AA.
 XX ADL71528;
 AC ADL71528;
 DT 20-MAY-2004 (first entry)
 XX

Novel human secreted protein seqid 132.

DE antiinflammatory; neuroprotective; nocrotropic; antiparkinsonian;
 KW anticonvulsant; antiinflammatory; CNS; gynaecological; antiarthritic;
 KW antiashmatic; anti-HIV; virocidic; endocrine; cytostatic;
 KW immunosuppressive; antiallergic; cardiovascular; respiratory;

KW dermatological; antimicrobial; gastrointestinal; gene therapy;
KW neurodegenerative disease; behavioral disorder; inflammatory condition;
KW hyperproliferative disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; metabolic disorder; Tay-Sachs disease;
KW Leash-Nyhan syndrome; reproductive disorder; immunological disorder;
KW arthritis; asthma; AIDS; endocrine disorder; immune disorder;
KW Hodgkin's lymphoma; haematopoietic disorder; muscular disorder;
KW Leukemia; autoimmune disorder; allergy; cancer; cardiovascular disorder;
KW respiratory disorder; pulmonary disorder; connective tissue disorder;
KW skin disorder; CNS disorder; congenital disorder; infectious disorder;
KW gastrointestinal disorder; human; secreted protein.
XX
XX Homo sapiens.
XX
XX US2004034196-A1.
XX
XX 19-FEB-2004.
XX
XX 27-JAN-2003; 2003US-00351334.
XX
XX 30-JUL-1998; 98US-0094657P.
XX 05-AUG-1998; 98US-0095486P.
XX 06-AUG-1998; 98US-0095454P.
XX 06-AUG-1998; 98US-0095455P.
XX 12-AUG-1998; 98US-0096319P.
XX 29-JUL-1999; 99WO-US017130.
XX 24-JAN-2000; 2000US-00489847.
XX 25-JAN-2002; 2002US-0350898P.
XX
XX (KOMA/) KOMATSOUIS G A.
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (DIJAN/) DIJAN D R.
XX (MOOR/) MOORE P A.
XX (SHIY/) SHI Y.
XX (LAFU/) LAFLEUR D W.
XX (WEIY/) WEI Y.
XX
XX Komatsoulis GA, Rosen CA, Ruben SM, Dian DR, Moore PA, Shi Y;
XX Lafleur DW, Wei Y;
XX
XX WPI; 2004-180094/17.
XX N-PSDB; ADL71412.
XX
XX New human secreted nucleic acid, useful for diagnosing and treating
XX neurodegenerative, inflammatory, hyperproliferative, metabolic,
XX reproductive, cardiovascular, respiratory or immunological disorders or
XX diseases.
XX
XX Claim 11; SEQ ID NO 132; 234bp; English.
XX
XX The invention describes an isolated human nucleic acid molecule (1)
XX comprising a polynucleotide having a nucleotide sequence at least 95%
XX identical to a sequence polynucleotide fragment of SEQ ID NO: X or of
XX the cDNA sequence included in ATCC Deposit No: Z, which is hybridisable
XX to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or
XX epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA
XX sequence included in ATCC Deposit No: Z, which is hybridisable to SEQ ID
XX NO: X, having a biological activity. The nucleic acids and polypeptides,
XX pharmaceutical formulations and kits are useful in diagnosing and
XX treating neurodegenerative diseases states, behavioral disorders,
XX inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's
XX disease, Parkinson's disease or Huntington's diseases), metabolic
XX disorders (e.g. Tay-Sachs disease or Leash-Nyhan syndrome), reproductive
XX disorders, immunological disorders (e.g. arthritis, asthma or AIDS),
XX endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic
XX or muscular disorders (e.g. leukemia), autoimmune disorders, allergy,
XX cancer, cardiovascular, respiratory or pulmonary disorders, disorders or
XX conditions afflicting connective tissue, skin disorders, CNS disorders,
XX congenital disorders, infectious disorders and gastrointestinal
XX disorders. This is the amino acid sequence of a novel human secreted
XX protein of the invention. Note: This sequence does not appear in the
XX printed specification but is available in electronic format from the US

CC patent office at ftp.segdata.uspro.gov/segdata.html?docID=20040034196.
XX
XX Sequence 709 AA;
XX
XX Query Match 82.8%; Score 3625.5; DB 8; Length 709;
XX Best Local Similarity 98.3%; Pred. No. 1.9e-308;
XX Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
XX
XX 1 MAGNCGVSIALLGVLLGAAALPRGAAFEIALPRESNITVLTKGPTLLAKPCYVI 60
XX 1 MAGNCGVSIALLGVLLGAAALPRGAAFEIALPRESNITVLTKGPTLLAKPCYVI 60
XX
XX 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQNIDCMGSPCFGEVQLOPSTSLPT 120
XX 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQNIDCMGSPCFGEVQLOPSTSLPT 120
XX
XX 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGGSCPDGVTHSISGRIDATVVRIGTFCSN 180
XX 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGGSCPDGVTHSISGRIDATVVRIGTFCSN 180
XX
XX 181 GTVSRITKQEGVKALHPMFHPRVSGFSTANSSIRLCITIEVFEGESATLMSANY 240
XX 181 GTVSRITKMXEGVKALHPMFHPRVSGFSTANSSIRLCITIEVFEGESATLMSANY 240
XX
XX 241 PEGFPEDELMTQFVPAHLBAVSFLNPNLSNCKRKEERYEYIPGSTTPEVFKLBDX 300
XX 241 PEGFPEDELMTQFVPAHLBAVSFLNPNLSNCKRKEERYEYIPGSTTPEVFKLBDX 300
XX
XX 301 QPGNMGFNLISLQCDQDAQSPGILRLQFOVLVQHPONESNKTIYVDLSNERASLTIE 360
XX 301 QPGNMGFNLISLQCDQDAQSPGILRLQFOVLVQHPONESNKTIYVDLSNERASLTIE 360
XX
XX 361 PRPVQSKRFVPGCVCSRTCSNLTLTSGSHKISFLDDTLTRLMNTEKITSCTDH 420
XX 361 PRPVQSKRFVPGCVCSRTCSNLTLTSGSHKISFLDDTLTRLMNTEKITSCTDH 420
XX
XX 421 RYCQKRSYSLQVSPDILHLYELHDFSWKLVPKDRSLVLPVQKLOQHTHEKPCNTSF 480
XX 421 RYCQKRSYSLQVSPDILHLYELHDFSWKLVPKDRSLVLPVQKLOQHTHEKPCNTSF 480
XX
XX 481 SYLVASAIPODLYFGSFCPGSITQIOVKONISVTLTFPAPSKOBSRQGLTVSFLPY 540
XX 481 SYLVASAIPODLYFGSFCPGSITQIOVKONISVTLTFPAPSKOBSRQGLTVSFLPY 540
XX
XX 541 FKEBGFVFTPDTSKVLRTPNMDRGPLSTLSVSMNIVRDOVACLTFFKERSGVVCO 600
XX 541 FKEBGFVFTPDTSKVLRTPNMDRGPLSTLSVSMNIVRDOVACLTFFKERSGVVCO 600
XX
XX 601 TGRAFMIIOBORTRAEELFSLDEDLPPKPSFHHSFWNISNCSPTSGKQDLFFSVTLT 660
XX 601 TGRAFMIIOBORTRAEELFSLDEDLPPKPSFHHSFWNISNCSPTSGKQDLFFSVTLT 660
XX
XX 661 PRVDLVILIAANGGVLLSALGLICVYKKKKTKNKBPANCIYXN 710
XX 661 PRVDLVILIAANGGVLLSALGLICVYKKKKTKNKBPANCIYXN 710
XX
XX 710 PRVDLVILIAANGGVLLSALGLICVYKKKKTKNKBPANCIYXN 710
XX 710 PRVDLVILIAANGGVLLSALGLICVYKKKKTKNKBPANCIYXN 710
XX
XX RESULT 7
XX AAM49640
XX ID AAM49640 standard; protein; 749 AA.
XX
XX AAM49640;
XX
XX 17-MAY-2002 (first entry)
XX
XX Human tumour-associated antigen B345 protein.
XX
XX Tumour-associated antigen; human; B345; cytosolic; cell communication;
XX cell interaction; signal transduction; metastasis; cancer; colon;
XX immunotherapy; carcinoma; lung; diagnosis.
XX
XX Homo sapiens.
XX

PN WO200204508-A1.
 XX
 XX 17-JAN-2002.
 XX
 XX 05-JUL-2001; 2001WO-EP007705.
 XX
 XX 07-JUL-2000; 2000DE-01033080.
 PR 19-APR-2001; 2001DE-01019294.
 XX
 PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 XX
 PI Schaefer N, Scherl-Mostageer M, Sommergruber W, Abseher R;
 XX
 DR MPI, 2002-171704/42.
 XX N-PSDB; ABA95506.
 PT New tumor-associated antigen B345, useful for diagnosis and immunotherapy
 PT of tumors, also related nucleic acid and antibodies.
 XX
 PS Example 6; Page 76-78; 102pp; German.
 XX
 CC This invention describes a novel tumor-associated antigen, designated
 CC B345 which has cytostatic activity. B345 is involved in communication,
 CC interaction and/or signal transduction with extracellular components and
 CC ligands, especially in the metastatic potential of cancers, particularly
 CC of the colon. B345 or its immunogenic fragments, also the DNA that
 CC encodes it, are useful for immunotherapy of cancer, particularly
 CC carcinoma of lung or colon. Antibodies raised against B345 are useful for
 CC treatment and diagnosis of cancers that are associated with B345
 CC expression, including their use for targeted delivery of cytotoxic or
 CC radioactive agents. Probes derived from B345 can be used to detect tumor
 CC B345 specific mutations in the B345 sequence, and can be used to screen for
 CC associated antigen modulators. This sequence represents a human B345 tumour-
 CC associated antigen described in the invention
 CC
 SQ Sequence 749 AA;
 Query Match 78.9%; Score 3454; DB 5; Length 749;
 Best Local Similarity 90.0%; Pred. No. 2,3e-293;
 Matches 674; Conservative 9; Mismatches 30; Indels 36; Gaps 6;
 QY 67 MLSTKSGRIYFTSCQSPENHFVIEIQKIDCMGSPCPGFEVOPQSTSLPTLNRTFI 126
 DB 1 MLSTKSGRIYFTSCQSPENHFVIEIQKIDCMGSPCPGFEVOPQSTSLPTLNRTFI 60
 QY 127 WDVAHAKSIGLELOFSIRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCNCTVSRI 186
 DB 61 WDVAHAKSIGLELOFSIRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCNCTVSRI 120
 QY 187 KMOGKVAKALHPMHPNNVSGFSIANRSSIKRLCIITSVEGESATLMANYPEGPE 246
 DB 121 KMOGKVAKALHPMHPNNVSGFSIANRSSIKRLCIITSVEGESATLMANYPEGPE 180
 QY 247 DELTMOQFVPAHRAVASFLNPLNSCERKEERYEYIPSTTNPEVFKLEDKOPGMA 306
 DB 181 DELTMOQFVPAHRAVASFLNPLNSCERKEERYEYIPSTTNPEVFKLEDKOPGMA 240
 QY 307 GNFPLSLQGCDDQADQSPGILRLQFQVLVQHPQNESNKIYVVDLSNERAMSITIEBRPVQ 366
 DB 241 GNFPLSLQGCDDQADQSPGILRLQFQVLVQHPQNESNKIYVVDLSNERAMSITIEBRPVQ 300
 QY 367 SRKVPQCGFVLEERTCSNLTITSGSKKISFLCDDLITRLMMNVEKTIISCTDHRVCOR 426
 DB 301 SRKVPQCGFVLEERTCSNLTITSGSKKISFLCDDLITRLMMNVEKTIISCTDHRVCOR 360
 QY 427 SYSLOVPSDILHLPELHDFSKWKLVPKDRSLVLPVPAKQLOQTHREPCNTSFSYLVA 486
 DB 361 SYSLOVPSDILHLPELHDFSKWKLVPKDRSLVLPVPAKQLOQTHREPCNTSFSYLVA 420
 QY 487 AIPQODLVFSGFCPGSGIKQIQVKNISVTLRTFAPSFXQASRQGLTVSFIPIYFKERG 546
 DB 421 AIPQODLVFSGFCPGSGIKQIQVKNISVTLRTFAPSFXQASRQGLTVSFIPIYFKERG 480

QY 547 FTVPTDTSKYVLRTPMNRDGLPSLTSYWNISVPRODVACTLPKERSGVCCGRFAM 606
 DB 481 FTVPTDTSKYVLRTPMNRDGLPSLTSYWNISVPRODVACTLPKERSGVCCGRFAM 540
 QY 607 IIOEQRTAEERIFSLDEVDLPKPSFHHHSFVWVINSNCSPTSGKOLDLFSVTLTPRTVDL 666
 DB 541 IIOEQRTAEERIFSLDEVDLPKPSFHHHSFVWVINSNCSPTSGKOLDLFSVTLTPRTVDL 600
 QY 667 TVILIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGIYXNINTEMPROPKKEQKGR 726
 DB 601 TVILIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGIYXNINTEMPROPKKEQKGR 659
 QY 727 KDNDSHYVAIVEDPMVYHLLQDSSGSLQPEVDITRFQGTMCV---PSP----- 776
 DB 660 -----ERTWT--PMCMQSSRT-----PMW--MGICVRIIPAAPSCSGRWT 694
 QY 777 PTICRAPTAKTATEEPPRSPPESEEP 805
 DB 695 PTICRAPTAKTATEEPPRSPPESEEP 723
 RESULT 8
 AAB95558
 ID AAB95558 standard; protein; 649 AA.
 XX
 AC AAB95558;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:18191.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR MPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 18191; 2537bp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX Sequence 649 AA;

XX Query Match 77.6%; Score 3395; DB 4; Length 649;

XX Best Local Similarity 99.2%; Pred. No. 2.8e-288;

XX Matches 644; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 188 MOEGVKMALHLPFPRNVSGFSINRSGIKRLCTIESVEEGSGATLMSANYPEGFPED 247

DB 1 MOEGVKMALHLPFPRNVSGFSINRSGIKRLCTIESVEEGSGATLMSANYPEGFPED 60

QY 248 ELMTQFVVPALRLASVSFLANPLNSCEKEREVEYITGTTNBPVFLDKKQGNMAG 307

DB 61 ELMTQFVVPALRLASVSFLANPLNSCEKEREVEYITGTTNBPVFLDKKQGNMAG 120

QY 308 NFNLISLOGDDDAOSPGILRLQFOYLVOHPONESNKIYVVDLSNBRAMSLTEPRPVKOS 367

DB 121 NFNLISLOGDDDAOSPGILRLQFOYLVOHPONESNKIYVVDLSNBRAMSLTEPRPVKOS 180

QY 368 RKFPVGCFCVCEISRTCSNLTLTSGSKHISFLCDLFTLMNVKETSICTDHRVCORKS 427

DB 181 RKFPVGCFCVCEISRTCSNLTLTSGSKHISFLCDLFTLMNVKETSICTDHRVCORKS 240

QY 428 YSLQVPSDILHLPVELHDPFWKLLVPKORLSLVVPAQQLQOHTHEKPNPNTSFLVAVSA 487

DB 241 YSLQVPSDILHLPVELHDPFWKLLVPKORLSLVVPAQQLQOHTHEKPNPNTSFLVAVSA 300

QY 488 IPSOOLYFGSPFCGGSIKOIOVKONISVLTFTAPSEFXOEASROGLTVSFIFYFKEGVF 547

DB 301 IPSOOLYFGSPFCGGSIKOIOVKONISVLTFTAPSEFXOEASROGLTVSFIFYFKEGVF 360

QY 548 TVTPTDKSKVYLRTPNMDRGLPSLTSSVSNISVPRDOVALCTFFPKERSGVVQOTGRAMI 607

DB 361 TVTPTDKSKVYLRTPNMDRGLPSLTSSVSNISVPRDOVALCTFFPKERSGVVQOTGRAMI 420

QY 608 IOBORTRAEITSLDEEDVLPKPSFHHHSFWNVISNCSPTSGQDLILFSTVLTPTVDLT 667

DB 421 IOBORTRAEITSLDEEDVLPKPSFHHHSFWNVISNCSPTSGQDLILFSTVLTPTVDLT 480

QY 668 VLLIAVGGGVLLLSALGLIICVKKKKKKTKNGPAVGIYNNININTEMROPKKFOKGRK 727

DB 481 VLLIAVGGGVLLLSALGLIICVKKKKKKTKNGPAVGIYNNININTEMROPKKFOKGRK 540

QY 728 DNDSHVYAVIEDTMYGHLLODSSGFLOPEVDYTRPFGGTGVCPPSPPTICSRAPTA 787

DB 541 DNDSHVYAVIEDTMYGHLLODSSGFLOPEVDYTRPFGGTGVCPPSPPTICSRAPTA 600

QY 788 LATEPPRPSPESESEPTTSHPNNGDVSSKDTDILPLKTQEPMEPAE 836

DB 601 LATEPPRPSPESESEPTTSHPNNGDVSSKDTDILPLKTQEPMEPAE 649

XX RESULT 9

XX ADL71669 standard; protein; 414 AA.

XX ADL71669;

XX 20-MAY-2004 (first entry)

XX Novel human secreted protein fragment seqid 273.

XX antiinflammatory; neuroprotective; nootropic; antiParkinsonian;

KW anticonvulsant; antilipemic; CNS; gynaecological; antiarthritic;
KW antiaesthetic; anti-HIV; virucide; endocrine; cytostatic;
KW immunosuppressive; anti-allergic; cardiovascular; respiratory;
KW dermatological; antimicrobial; gastrointestinal; gene therapy;
KW neurodegenerative disease; behavioral disorder; inflammatory condition;
KW hyperproliferative disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; metabolic disorder; Tay-Sach's disease;
KW Leash-Nyhan syndrome; reproductive disorder; immunological disorder;
KW arthritis; asthma; AIDS; endocrine disorder; immune disorder;
KW Hodgkin's lymphoma; haematopoietic disorder; muscular disorder;
KW leukaemia; autoimmune disorder; allergy; cancer; cardiovascular disorder;
KW respiratory disorder; pulmonary disorder; connective tissue disorder;
KW skin disorder; CNS disorder; congenital disorder; infectious disorder;
KW gastrointestinal disorder; human; secreted protein.

OS Homo sapiens.

PN US2004034196-A1.

PD 19-FEB-2004.

PF 27-JAN-2003; 2003US-00351334.

PR 30-JUL-1998; 98US-0094657P.

PR 05-AUG-1998; 98US-0095486P.

PR 06-AUG-1998; 98US-0095454P.

PR 06-AUG-1998; 98US-0095455P.

PR 12-AUG-1998; 98US-0096319P.

PR 29-JUL-1999; 99WO-US017130.

PR 24-JAN-2000; 2000US-00489847.

PR 25-JAN-2002; 2002US-0350898P.

PA (KOMA/) KOMATSU S. I. G. A.

PA (ROSE/) ROSEN C. A.

PA (RUBE/) RUBEN S. M.

PA (DUAN/) DUAN D. R.

PA (MOOR/) MOORE P. A.

PA (SHIY/) SHI Y.

PA (LAFLE/) LAFLEUR D. W.

PA (WEIY/) WEI Y.

PI Komatsoulis GA, Rosen CA, Ruben SM, Duan DR, Moore PA, Shi Y;

PI Lafleur DW, Wei Y;

PI MPI, 2004-180094/17.

DR MPI, 2004-180094/17.

XX New human secreted nucleic acid, useful for diagnosing and treating

XX neurodegenerative, inflammatory, hyperproliferative, metabolic,

XX reproductive, cardiovascular, respiratory or immunological disorders or

XX diseases.

PS Disclosure: SEQ ID NO 273; 234pp; English.

The invention describes an isolated human nucleic acid molecule (1)
comprising a polynucleotide having a nucleotide sequence at least 95%
identical to a sequence polynucleotide fragment of SEQ ID NO: X or of
the cDNA sequence included in ATCC Deposit No: Z, which is hybridizable
to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or
epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA
sequence included in ATCC Deposit No: Z, which is hybridizable to SEQ ID
NO: X, having a biological activity. The nucleic acids and polypeptides,
pharmaceutical formulations and kits are useful in diagnosing and
treating neurodegenerative diseases states, behavioral disorders,
inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's
disease, Parkinson's disease or Huntington's disease), metabolic
disorders (e.g. Tay-Sach's disease or Leash-Nyhan syndrome), reproductive
disorders, immunological disorders (e.g. arthritis, asthma or AIDS),
endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic
or muscular disorders (e.g. leukaemia), autoimmune disorders, allergy,
cancer, cardiovascular, respiratory or pulmonary disorders, disorders or
conditions afflicting connective tissue, skin disorders, CNS disorders,
congenital disorders, infectious disorders and gastrointestinal
disorders. This is the amino acid sequence of a novel human secreted

CC Protein fragment of the invention. Note: This sequence does not appear in
 CC the printed specification but is available in electronic format from the
 CC US patent office at ftp.segdata.uspto.gov/segdata.html?DocID=20040034196.

XX Sequence 414 AA;

Query Match 49.2%; Score 2155; DB 8; Length 414;
 Best Local Similarity 100.0%; Pred. No. 9,4e-180;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPKGAFAFIALPRESNTIVLIKGTPTLAKPCYIV 60
 DB 1 MAGNCGVSIALGLVLLGAARLPKGAFAFIALPRESNTIVLIKGTPTLAKPCYIV 60
 QY 61 SKRHITMLSTKSGRIVFTFSCQSPENHFVIEIQKNDMSGPCFGEVQLQPSLSLPT 120
 DB 61 SKRHITMLSTKSGRIVFTFSCQSPENHFVIEIQKNDMSGPCFGEVQLQPSLSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLELQFSIRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQFSIRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANY 240
 DB 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANY 240
 QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEBEVEYIIPGSTTNEVEFKLBDK 300
 DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEBEVEYIIPGSTTNEVEFKLBDK 300
 QY 301 QPGMAGNFNLSTGCGDDAQSFGILRLQFVLVQHPQNESKITYVDLSNERAMSLTTE 360
 DB 301 QPGMAGNFNLSTGCGDDAQSFGILRLQFVLVQHPQNESKITYVDLSNERAMSLTTE 360
 QY 361 PRPVQSRKFPVPGCFVCLSESTCSNLTLSGSKKISFLCDDLTRLMMANVEK 413
 DB 361 PRPVQSRKFPVPGCFVCLSESTCSNLTLSGSKKISFLCDDLTRLMMANVEK 413

RESULT 10
 AAY91592
 ID AAY91592 standard; protein; 443 AA.

XX AAY91592;

DT 29-JUN-2000 (first entry)

DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:265.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KM antiHIV; antiinflammatory; nootropic; neuroprotective; antiasthmatic;
 KM osteopathic; antidiabetic; antibacterial; anticancer; antidiabetic; antidiabetic;
 KM antipruritic; cardiac; gene therapy; cancer; neurological disorder;
 XX immune disease; inflammation; blood disorder; tumour.

OS Homo sapiens.

XX WO200006698-A1.

PD 10-FEB-2000.

PF 29-JUL-1999; 99WO-US017130.

PR 30-JUL-1998; 98US-0094657P.

PR 05-AUG-1998; 98US-0095486P.

PR 06-AUG-1998; 98US-0095455P.

PR 12-AUG-1998; 98US-0096319P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y,
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;

PI Soppet DR, Endress GA, Edner R, Olsen HS, Mucenaki M;

XX WPI; 2000-195282/17.

PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancer, neurological
 PT disorders, immune diseases, inflammation or blood disorders.

PS Disclosure; Page 18; 634pp; English.

XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;
 CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
 CC antiallergic; osteopathic; antidiabetic; antibacterial; antidiabetic;
 CC antidiabetic; antipruritic; cardiac; gene therapy; cancer; neurological disorder;
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
 CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
 CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC reproductive disorders, gastrointestinal disorders, respiratory disorders
 CC and metabolic disorders. The proteins or polynucleotides can also be used
 CC as food additives or preservatives. The proteins are also useful for
 CC identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are
 CC sequences used in the exemplification of the present invention

XX Sequence 443 AA;

Query Match 49.2%; Score 2155; DB 3; Length 443;
 Best Local Similarity 100.0%; Pred. No. 1e-179;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPKGAFAFIALPRESNTIVLIKGTPTLAKPCYIV 60
 DB 30 MAGNCGVSIALGLVLLGAARLPKGAFAFIALPRESNTIVLIKGTPTLAKPCYIV 89
 QY 61 SKRHITMLSTKSGRIVFTFSCQSPENHFVIEIQKNDMSGPCFGEVQLQPSLSLPT 120
 DB 90 SKRHITMLSTKSGRIVFTFSCQSPENHFVIEIQKNDMSGPCFGEVQLQPSLSLPT 149
 QY 121 LNRFTIMDVKAHKSIGLELQFSIRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 150 LNRFTIMDVKAHKSIGLELQFSIRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 209
 QY 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANY 240
 DB 210 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANY 269
 QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEBEVEYIIPGSTTNEVEFKLBDK 300
 DB 270 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEBEVEYIIPGSTTNEVEFKLBDK 329
 QY 301 QPGMAGNFNLSTGCGDDAQSFGILRLQFVLVQHPQNESKITYVDLSNERAMSLTTE 360
 DB 330 QPGMAGNFNLSTGCGDDAQSFGILRLQFVLVQHPQNESKITYVDLSNERAMSLTTE 389
 QY 361 PRPVQSRKFPVPGCFVCLSESTCSNLTLSGSKKISFLCDDLTRLMMANVEK 413
 DB 390 PRPVQSRKFPVPGCFVCLSESTCSNLTLSGSKKISFLCDDLTRLMMANVEK 442

RESULT 11
 ADL71667

ID ADL71667 standard; protein; 443 AA.
 AC ADL71667;
 XX
 XX 20-MAY-2004 (first entry)
 DT
 DE Novel human secreted protein fragment seqid 271.
 KM antiinflammatory; neuroprotective; nootropic; antiparkinsonian;
 KM anticonvulsant; antilipemic; CNS; gynaecological; antiarthritic;
 KM antitachycardic; anti-HIV; virucide; endocrine; cytostatic;
 KM immunosuppressive; antiallergic; cardiovascular; respiratory;
 KM dermatological; antimicrobial; gastrointestinal; gene therapy;
 KM neurodegenerative disease; behavioral disorder; inflammatory condition;
 KM hyperproliferative disorder; Alzheimer's disease; Parkinson's disease;
 KM Huntington's disease; metabolic disorder; Tay-Sachs disease;
 KM leaah-Nyhan syndrome; reproductive disorder; immunological disorder;
 KM arthritic; asthma; AIDS; endocrine disorder; immune disorder;
 KM Hodgkin's lymphoma; haematopoietic disorder; muscular disorder;
 KM leukaemia; autoimmune disorder; allergy; cancer; cardiovascular disorder;
 KM respiratory disorder; pulmonary disorder; connective tissue disorder;
 KM skin disorder; CNS disorder; congenital disorder; infectious disorder;
 KM gastrointestinal disorder; human; secreted protein.
 OS Homo sapiens.
 XX
 XX US2004034196-A1.
 PD 19-FEB-2004.
 XX
 PF 27-JAN-2003; 2003US-00351334.
 XX
 PR 30-JUN-1998; 98US-0094657P.
 PR 05-AUG-1998; 98US-0095486P.
 PR 06-AUG-1998; 98US-0095454P.
 PR 12-AUG-1998; 98US-0095455P.
 PR 29-JUN-1999; 99US-0096319P.
 PR 24-JAN-2000; 2000US-00489847.
 PR 25-JAN-2002; 2002US-0350898P.
 XX
 PA (KOMA/) KOMATSOUJIS G A.
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (DUAN/) DUAN D R.
 PA (MOORE/) MOORE P A.
 PA (SHIY/) SHI Y.
 PA (LAFU/) LAFLEUR D W.
 PA (WEIY/) WEI Y.
 XX
 PI Komatsoulis GA, Rosen CA, Ruben SM, Duan DR, Moore PA, Shi Y;
 PI Lafleur DW, Wei Y;
 XX
 DR WPI; 2004-180094/17.
 XX
 PT New human secreted nucleic acid, useful for diagnosing and treating
 PT neurodegenerative, inflammatory, hyperproliferative, metabolic,
 PT reproductive, cardiovascular, respiratory or immunological disorders or
 PT diseases.
 PS
 PS Disclosure; SEQ ID NO 271; 234pp; English.
 XX
 CC The invention describes an isolated human nucleic acid molecule (I)
 CC comprising a polynucleotide having a nucleotide sequence at least 95%
 CC identical to: a sequence polynucleotide fragment of SEQ ID NO: X or of
 CC the cDNA sequence included in ATCC Deposit No: Z, which is hybridisable
 CC to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or
 CC epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA
 CC sequence included in ATCC Deposit No: Z, which is hybridisable to SEQ ID
 CC NO: X, having a biological activity. The nucleic acids and polypeptides,
 CC pharmaceutical formulations and kits are useful in diagnosing and
 CC treating neurodegenerative diseases states, behavioral disorders,
 CC inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's

CC disease, Parkinson's diseases or Huntington's diseases), metabolic
 CC disorders (e.g. Tay-Sachs's disease or leaah-Nyhan syndrome), reproductive
 CC disorders, immunological disorders (e.g. arthritis, asthma or AIDS),
 CC endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic
 CC or muscular disorders (e.g. leukaemia), autoimmune disorders, allergy,
 CC cancer, cardiovascular, respiratory or pulmonary disorders, disorders or
 CC conditions affecting connective tissue, skin disorders, CNS disorders,
 CC congenital disorders, infectious disorders and gastrointestinal
 CC disorders. This is the amino acid sequence of a novel human secreted
 CC protein fragment of the invention. Note: This sequence does not appear in
 CC the printed specification but is available in electronic format from the
 CC US patent office at ftp.seqdata.uspo.gov/seqdata.html?docid=2004034196.
 XX
 SQ Sequence 443 AA;
 Query Match 49.2%; Score 2155; DB 8; Length 443;
 Best Local Similarity 100.0%; Pred. No. 1e-179;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGINCQSVIALGVLLGAAARLPRAAEAFIALPRESNTIVLTKGPTLLAKPCYVI 60
 DB 30 MAGINCQSVIALGVLLGAAARLPRAAEAFIALPRESNTIVLTKGPTLLAKPCYVI 89
 QY 61 SKRHITMLSIKSGRIYFTFSCQSPENHFVIEIQNIDCMGPGCFGEVQLQPSLSLPT 120
 DB 90 SKRHITMLSIKSGRIYFTFSCQSPENHFVIEIQNIDCMGPGCFGEVQLQPSLSLPT 149
 QY 121 LNRFTIWDVKAHKSIGLEQFSIRLROIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
 DB 150 LNRFTIWDVKAHKSIGLEQFSIRLROIGGESCPCDVTHSISGRIDATVVRIGTFCSN 209
 QY 181 GTVSRIKMQEGVKALHLPMFHPRNVSQFSIANSSIRKLCITISVFEGESATLMSANY 240
 DB 210 GTVSRIKMQEGVKALHLPMFHPRNVSQFSIANSSIRKLCITISVFEGESATLMSANY 269
 QY 241 PEGPEDELMTQFVPAHLRAVSFLNFNLSNCRKEREVEYIIPGSTTPEVFKLEDK 300
 DB 270 PEGPEDELMTQFVPAHLRAVSFLNFNLSNCRKEREVEYIIPGSTTPEVFKLEDK 329
 QY 301 QPGMAGNPNLSLQCCDDAOSPGILRLQFOVLVQHPNENSKYIYVDLSNERAMSLTIE 360
 DB 330 QPGMAGNPNLSLQCCDDAOSPGILRLQFOVLVQHPNENSKYIYVDLSNERAMSLTIE 389
 QY 361 PRPVKQSRKVPFGCFVCLSESTCSNLTLTSGSKHKSIFLCDLTRLMMNVEK 413
 DB 390 PRPVKQSRKVPFGCFVCLSESTCSNLTLTSGSKHKSIFLCDLTRLMMNVEK 442
 RESULT 12
 ID AAU12252
 XX AAU12252 standard; protein; 343 AA.
 AC AAU12252;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human PRO5773 polypeptide sequence.
 XX
 KM Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
 KM prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
 KM ear; proliferation; glucose; free fatty acid; skeletal acid; adipocyte;
 KM A-peptide; factor VIIA; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200140466-A2.
 XX
 PD 07-JUN-2001.
 PF 01-DEC-2000; 2000WO-US032678.
 XX
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.

PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028554.
 PR 09-DEC-1999; 99WO-US028565.
 PR 09-DEC-1999; 99US-0170262P.
 PR 16-DEC-1999; 99WO-US030095S.
 PR 20-DEC-1999; 99WO-US0300911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 06-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 03-MAR-2000; 2000US-0187202P.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209833P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 XX
 PA (GERTH) GENENTECH INC.
 XX
 PI Baker KP, Barresi M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerlesen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI; 2001-408281/43.
 DR N-PSDB; AAS21324.
 XX
 PT Isolated , secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing PRO
 PT polypeptides, and detect the presence of mammalian tumors e.g. lung,
 PT breast, prostate, cervical.
 XX
 PS Claim 12; Fig 162; 813pp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane PRO
 CC polypeptides. The PRO polypeptides are useful to detect other PRO
 CC polypeptides, to link bioactive molecules to cells expressing PRO
 CC polypeptides, to modulate biological activities of cells expressing PRO
 CC polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumors by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample. Some
 CC of the 275 sequences are also useful to stimulate the release of tumour
 CC neurosis factor-alpha (TNF-alpha) from human blood, the proliferation or
 CC differentiation of chondrocytes, the proliferation or gene expression in
 CC pericyte cells, the release of proteoglycans from cartilage, the
 CC proliferation of inner ear utricular supporting cells or of T-
 CC lymphocytes, the release of a cytokine from peripheral blood monocytes
 CC (pBMCs), or the proliferation of endothelial cells. Some of the PRO
 CC polypeptides may modulate glucose or free fatty acid uptake by skeletal
 CC muscle cells or by adipocytes, or inhibit binding of A-peptide to factor
 CC VIIA. The PRO polypeptides can be used in assays to identify molecules
 CC involved in binding interactions. The polynucleotides encoding PRO
 CC polypeptides can be used to generate probes, antisense RNA/DNA,

CC transgenic or knock out animals and can be used in gene therapy
 XX
 SQ Sequence 343 AA;
 Query Match 40.6%; Score 1779; DB 4; Length 343;
 Best Local Similarity 99.4%; Pred. No. 7e-147;
 Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGLNCGVSIALLGVLLGAAALPRGADEFIALPRESNITVLKLGPTLLAKPCYIVI 60
 DB 1 MAGLNCGVSIALLGVLLGAAALPRGADEFIALPRESNITVLKLGPTLLAKPCYIVI 60
 QY 61 SKRIITMLSTKSGRIYFTSCSPENHFVIEIKNIDCMGSPCPFGVQLQPSLSLPT 120
 DB 61 SKRIITMLSTKSGRIYFTSCSPENHFVIEIKNIDCMGSPCPFGVQLQPSLSLPT 120
 QY 121 LNRFTFMDVKAKHSIGLELOPSIRLRIOIGGSCSPGVTSISGRIDATVVRIGTFPCSN 180
 DB 121 LNRFTFMDVKAKHSIGLELOPSIRLRIOIGGSCSPGVTSISGRIDATVVRIGTFPCSN 180
 QY 181 GTVSRIRKQEGVKXALHPFHPRNVSQFSIANBSIKRLCIIISVFEGBGSATLMSANY 240
 DB 181 GTVSRIRKQEGVKXALHPFHPRNVSQFSIANBSIKRLCIIISVFEGBGSATLMSANY 240
 QY 241 PEGFPEDELMTWQVVPVAHLRASVSLNPNLSNCRKEERYEYIIPGSTTPPEYFKLEDK 300
 DB 241 PEGFPEDELMTWQVVPVAHLRASVSLNPNLSNCRKEERYEYIIPGSTTPPEYFKLEDK 300
 QY 301 QPGMAGNFNLISLGGCDODASPGIILQOVLVQHPONESNK 343
 DB 301 QPGMAGNFNLISLGGCDODASPGIILQOVLVQHPONESSE 343
 RESULT 13
 ABO17696
 ID ABO17696 standard; protein; 343 AA.
 XX
 AC ABO17696;
 XX
 DT 26-AUG-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO5773.
 XX
 KW Human, secreted and transmembrane protein; PRO; antiinflammatory;
 KW antiarteriosclerotic; cardiatic; anti-infertility; anti-HIV; cytostatic;
 KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
 KW TNF-alpha release; cell proliferation; cell differentiation;
 KW gene expression modulator; proteoglycan release; cytokine release;
 KW tumour; inflammatory disease; organ failure; atherosclerosis;
 KW cardiac injury; infertility; birth defect; premature aging; AIDS;
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;
 KW chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
 KW bioreactor; tissue typing.
 XX
 OS Homo sapiens.
 XX
 PN US2003032156-A1.
 XX
 PD 13-FEB-2003.
 XX
 PF 06-MAY-2002; 2002US-00140474.
 XX
 XX 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019093.
 PR 14-SEP-1998; 98WO-US019094.
 PR 14-SEP-1998; 98WO-US019177.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005130.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010723.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021030.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028554.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007337.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747255.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.

PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.

PA (GETH) GENENTECH INC.
XX
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AB, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2003-341980/32.
DR N-PSDB; ACD23933.
XX
XX
PT New secreted and transmembrane PRO nucleic acids, for treating
PT inflammation, organ failure, atherosclerosis, cardiac injury,
PT infertility, birth defects, premature aging, acquired immunodeficiency
PT syndrome (AIDS), or cancer.
XX
XX
PS Claim 12; Fig 162; 660pp; English.
XX
XX

The invention describes an isolated nucleic acid (I) comprising, or which has 80 % sequence identity to, or the full-length coding sequence of, one of 275 nucleotide sequences, and which encodes a corresponding CC polypeptide selected from 275 amino acid sequences, where all sequences CC are given in the specification. The polypeptide encoded by (I) is used to CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a CC PRO polypeptide, modulate a biological activity of a cell, stimulate the CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit CC the proliferation or differentiation of cells or gene expression, CC stimulate the release of proteoglycans, stimulate the release of cytokine CC from peripheral blood mononuclear cells, inhibit the binding of A-betaide CC to factor VIIa, or detect the presence of tumour in a mammal. The nucleic CC acid and polypeptide encoded by it, are useful for treating inflammatory CC diseases, organ failure, atherosclerosis, cardiac injury, infertility, CC birth defects, premature aging, acquired immunodeficiency syndrome CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as CC hybridisation probes, in chromosome and gene mapping, and in generating CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals, CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing. CC This is the amino acid sequence of a novel human secreted and CC transmembrane PRO polypeptide

XX Sequence 343 AA;

Query Match 40.6%; Score 1779; DB 6; Length 343;

Best Local Similarity 99.4%; Pred. No. 7e-147; Indels 0; Gaps 0;

Matches 341; Conservative 2; Mismatches 0;

QY 1 MAGLNCGVSIALLGTLGAAALPRGAEAFIALPRESNITVLTKGPTLLAKRCYVI 60
DB 1 MAGLNCGVSIALLGTLGAAALPRGAEAFIALPRESNITVLTKGPTLLAKRCYVI 60
QY 61 SKRHITMLSIKSGRIIVTFSCSPENHFVIEIQKNDICMGCPDPFGEVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGRIIVTFSCSPENHFVIEIQKNDICMGCPDPFGEVQLOPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLELQFSIPRLROIIGSGESGPODVTISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELQFSIPRLROIIGSGESGPODVTISGRIDATVVRIGTFCSN 180

QY 181 GTVSRKQEGVKMLHLPMFPRVSGFSIANRSSIKRLCTIESVFEGGSATLMSANY 240
Db 181 GTVSRKQEGVKMLHLPMFPRVSGFSIANRSSIKRLCTIESVFEGGSATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALRLASVSFLNFINLSNCKEKEERYEYYPGSTTPEVFKLBDK 300
Db 241 PEGFPEDELMTQFVVPALRLASVSFLNFINLSNCKEKEERYEYYPGSTTPEVFKLBDK 300
QY 301 QPGMAGNFNLISLQCCDDAQSPGILRLQFVLVQHVPONESNK 343
Db 301 QPGMAGNFNLISLQCCDDAQSPGILRLQFVLVQHVPONESSE 343
RESULT 14
ID ABU80950 standard; protein, 343 AA.
XX
AC ABU80950;
XX
DT 23-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #81.
XX
KM Human; PRO polypeptide; secreted and transmembrane protein;
KM anti-PRO antibody; diagnostic assay; gene expression; diabetes;
KM bone disorder; cartilage disorder; rheumatoid arthritis; obesity;
KM sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;
KM hearing loss; coagulation disorder; stroke; heart attack; cardiacnt;
KM antidiabetic; anorectic; vulnerrary; antiarthritic; osteopathic;
KM antirheumatic; auditory; cerebroprotective; angiogenic.
XX
OS Homo sapiens.
XX
PN US200304311-A1.
XX
PD 02-JAN-2003.
XX
PF 19-DEC-2001; 2001US-00028072.
XX
PR 18-JUN-1997; 97US-0049911P.
PR 26-AUG-1997; 97US-0056974P.
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059122P.
PR 18-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 19-SEP-1997; 97US-0059352P.
PR 19-SEP-1997; 97US-0059588P.
PR 24-SEP-1997; 97US-0059836P.
PR 17-OCT-1997; 97US-0062250P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 17-OCT-1997; 97US-0063755P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063082P.
PR 24-OCT-1997; 97US-0063127P.
PR 27-OCT-1997; 97US-0063322P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063561P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063733P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.

PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 11-DEC-1997; 97US-0069212P.
PR 11-DEC-1997; 97US-0069278P.
PR 11-DEC-1997; 97US-0069334P.
PR 16-DEC-1997; 97US-0069694P.
PR 23-JAN-1998; 98US-0072320P.
PR 04-FEB-1998; 98US-0073612P.
PR 09-FEB-1998; 98US-0074086P.
PR 12-MAR-1998; 98US-0077791P.
PR 20-MAR-1998; 98US-0078910P.
PR 25-MAR-1998; 98US-0079294P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079728P.
PR 31-MAR-1998; 98US-0080165P.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028651.
PR 02-DEC-1999; 99WO-US028664.
PR 02-DEC-1999; 99WO-US028665.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030939.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US003376.
PR 18-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
XX
XX (GENTH) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Geritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

DR WPI: 2003-352836/33.
DR N-PSDB; ACA67074.
PT New isolated PRO polypeptide useful for treating diabetes, rheumatoid
PT arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or
PT heart attack.
XX
PS Claim 12; Fig 162; 643pp; English.
XX The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a medicament
CC useful in the treatment of diabetes, bone and/or cartilage disorders
CC (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,
CC hyper- or hypo-insulinemia, hearing loss, and coagulation disorders
CC (e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic
CC assays for PRO, by detecting its expression in specific cells, tissues or
CC serum, and for affinity purification of PRO from recombinant cell culture
CC or natural sources. AB060870-AB081144 represent the human PRO
CC polypeptides of the invention. Note: The sequence data for this patent
CC was obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/patident.htm
XX
SQ Sequence 343 AA;
Query Match 40.6%; Score 1779; DB 6; Length 343;
Best Local Similarity 99.4%; Pred. No. 76-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGLNCGSIALGLGALPGAAFEIALPRESNITVLIKGTPTLLAKPCYIV 60
DB 1 MAGLNCGSIALGLGALPGAAFEIALPRESNITVLIKGTPTLLAKCYIV 60
QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQNIDCMGCPGFEVQLQPTSLAPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQNIDCMGCPGFEVQLQPTSLAPT 120
QY 121 LNRFTIWKVKAHKSIGLELOPSIPRLROI GPESCPDGVTHSISRIDATVVRIGTFPSN 180
DB 121 LNRFTIWKVKAHKSIGLELOPSIPRLROI GPESCPDGVTHSISRIDATVVRIGTFPSN 180
QY 181 GTVSRIKMGQGVKMLHPWFHPRNVSGFSIANRSSIKRLCTIESEVFEGSGATLMSANY 240
DB 181 GTVSRIKMGQGVKMLHPWFHPRNVSGFSIANRSSIKRLCTIESEVFEGSGATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALHRAVSFLNPNLSNCRKEERVEYYIPGSTTNEVFLEDK 300
DB 241 PEGFPEDELMTWQFVVPALHRAVSFLNPNLSNCRKEERVEYYIPGSTTNEVFLEDK 300
QY 301 QPGNAGNPNLSLOCCDDAOSPGILRLOFQVLNQHPNESWK 343
DB 301 QPGNAGNPNLSLOCCDDAOSPGILRLOFQVLNQHPNESSE 343
RESULT 15
AB066650
ID AB066650 standard; protein; 343 AA.
XX AB066650;
AC
XX 23-MAY-2003 (first entry)
DT
XX Human PRO polypeptide #81.
DE
XX Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
XX differentiation; chondrocyte; tumour; genetic disorder; cytostatic.
OS Homo sapiens.
XX US2003036180-A1.
XX

PD 20-FEB-2003.
XX
PF 09-MAY-2002; 2002US-00143114.
XX
PR 31-MAR-1997; 97WO-US0052230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019053.
PR 14-SEP-1998; 98WO-US019094.
PR 16-SEP-1998; 98WO-US019177.
PR 17-SEP-1998; 98WO-US019330.
PR 17-OCT-1998; 98WO-US019437.
PR 29-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 20-NOV-1998; 98WO-US022992.
PR 01-DEC-1998; 98WO-US024855.
PR 05-JAN-1999; 98WO-US025108.
PR 08-MAR-1999; 99WO-US000106.
PR 10-MAR-1999; 99WO-US005028.
PR 20-APR-1999; 99WO-US005150.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020554.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028554.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 22-DEC-1999; 99WO-US030911.
PR 30-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 05-JAN-2000; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004514.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.

PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001US-00806666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 05-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001US-00874503.
PR 19-JUN-2001; 2001US-00882636.
PR 20-JUN-2001; 2001US-00886342.
PR 21-JUN-2001; 2001WO-US019692.
PR 22-JUN-2001; 2001US-00887879.
PR 29-JUN-2001; 2001WO-US020116.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
PI Gerltsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-332040/31.
XX N-PSDB; ACA03683.
XX
XX New secreted and transmembrane PRO nucleic acids, useful for gene
PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue
PT typing, and in chromosome identification.
XX
XX
XX Claim 12; Fig 162; 660pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides are useful for detecting other PRO polypeptides, for linking
CC bioactive molecules to cells expressing PRO polypeptides, for modulating
CC biological activities of cells expressing PRO polypeptides, and for
CC identifying agonists or antagonists. The PRO polypeptides are useful for
CC for stimulating the release of tumour necrosis factor (TNF)-alpha from
CC human blood, for stimulating the proliferation or differentiation of
CC chondrocytes, and detecting the presence of tumours. The polynucleotide
CC sequences encoding PRO polypeptides are useful as hybridisation probes,
CC in chromosome and gene mapping, in the generation of antisense RNA and
CC DNA, in the preparation of PRO polypeptides, for generating transgenic
CC animals or knockout animals, for the genetic analysis of individuals with
CC genetic disorders, and in gene therapy. ABU6570-ABU6844 represent the
CC human PRO polypeptides of the invention. Note: The sequence data for this
CC patent was obtained in electronic format directly from the USPRO web site
CC at seqdata.uspto.gov/patidbIdentify.html
XX
XX Sequence 343 AA:

SO

Query Match 40.6%; Score 1779; DB 6; Length 343;
Best Local Similarity 99.4%; Pred. No. 7e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCGVSTALLGVLLGGAARLRPGAAPEIALPRESNITVLKLGPTLLAKPCYIVI 60
|||||

Db 1 MAGINCGVSTALLGVLLGGAARLRPGAAPEIALPRESNITVLKLGPTLLAKPCYIVI 60
QY 61 SKRHITMLSTIKSGRIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQIQPSTSLPT 120
Db 61 SKRHITMLSTIKSGRIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQIQPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEQFSIPRLROIQGESCPGVYHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIWDVKAHKSIGLEQFSIPRLROIQGESCPGVYHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRITKMGEGVKMALHPMPHPNVSQFSIANKRSIKRLCTIEVFEGEGSATIMSANY 240
Db 181 GTVSRITKMGEGVKMALHPMPHPNVSQFSIANKRSIKRLCTIEVFEGEGSATIMSANY 240
QY 241 PEGFPEDELMTWQFVVPBAHLRASVSLNPNLSNCRKEERYEYIIPGTTNPEVFKLEDK 300
Db 241 PEGFPEDELMTWQFVVPBAHLRASVSLNPNLSNCRKEERYEYIIPGTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFVLYVQHQNESNK 343
Db 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFVLYVQHQNESSE 343

Search completed: May 4, 2006, 22:54:39
Job time : 138 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:25:30 ; Search time 47 Seconds
(without alignments)
1470.572 Million cell updates/sec

Title: US-10-781-564-1

Sequence: 1 MAGLNCVSTALLGVLLGA.....SSKDTDIPLKTQRPMPAP 836

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/6.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/H.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/RE.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiletest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3625.5	82.8	709	2	US-09-489-847-132
2	2155	49.2	414	2	US-09-489-847-273
3	2155	49.2	443	2	US-09-489-847-271
4	827	18.9	159	2	US-09-489-847-272
5	141	3.2	3623	2	US-09-341-461-2
6	126.5	2.9	590	2	US-09-902-540-11386
7	118.5	2.7	2476	1	US-08-276-967-2
8	118	2.7	2766	2	US-09-964-956-62
9	114	2.6	1589	2	US-08-755-587-189
10	112.5	2.6	440	2	US-09-538-092-999
11	112.5	2.6	451	2	US-09-949-016-9282
12	112.5	2.6	1694	2	US-09-560-385A-12
13	112.5	2.6	1725	2	US-09-560-385A-10
14	108	2.5	699	2	US-09-949-016-6138
15	107.5	2.5	649	2	US-09-248-796A-20641
16	107.5	2.5	744	2	US-09-854-845-43
17	107.5	2.5	771	2	US-09-854-845-47
18	107.5	2.5	843	2	US-09-854-845-27
19	107.5	2.5	870	2	US-09-854-845-31
20	107.5	2.5	1719	1	US-08-459-568-4
21	107.5	2.5	1719	1	US-08-399-411-4
22	107.5	2.5	1719	2	US-08-516-859A-4
23	107.5	2.5	1719	2	US-09-586-472-4
24	107.5	2.5	1719	2	US-09-528-706-4
25	107.5	2.5	1719	2	US-10-024-450-4
26	107	2.4	739	2	US-09-854-845-45
27	107	2.4	766	2	US-09-854-845-49

ALIGNMENTS

28	107	2.4	838	2	US-09-854-845-29	Sequence 29, Appl
29	107	2.4	865	2	US-09-854-845-33	Sequence 33, Appl
30	106.5	2.4	2343	2	US-09-324-867-2	Sequence 2, Appl
31	106	2.4	670	1	US-08-473-750-11	Sequence 11, Appl
32	106	2.4	670	1	US-08-477-326-11	Sequence 11, Appl
33	105.5	2.4	369	2	US-09-107-532A-5754	Sequence 5754, Ap
34	105.5	2.4	690	2	US-09-248-796A-19169	Sequence 19169, A
35	105	2.4	1540	2	US-09-949-016-11382	Sequence 11382, A
36	105	2.4	1540	2	US-09-949-016-11383	Sequence 11383, A
37	104	2.4	449	1	US-08-839-008-2	Sequence 2, Appl
38	104	2.4	458	2	US-09-949-016-7238	Sequence 7238, Ap
39	103.5	2.4	683	2	US-09-213-293D-1	Sequence 1, Appl
40	103.5	2.4	727	1	US-08-475-844-9	Sequence 9, Appl
41	103.5	2.4	727	4	PCT-US95-08429-9	Sequence 9, Appl
42	103.5	2.4	736	2	US-09-949-016-9241	Sequence 9241, Ap
43	103.5	2.4	866	2	US-09-527-084A-4	Sequence 4, Appl
44	103	2.4	449	1	US-08-839-008-9	Sequence 9, Appl
45	103	2.4	449	2	US-09-919-497-89	Sequence 89, Appl

RESULT 1

US-09-489-847-132

Sequence 132, Application US/09489847

Patent No. 6476195

GENERAL INFORMATION:

APPLICANT: Rosen et al

TITLE OF INVENTION: 98 Human Secreted Proteins

FILE REFERENCE: P2031P1

CURRENT FILING DATE: US/09/489, 847

EARLIER FILING DATE: 2000-01-24

EARLIER APPLICATION NUMBER: PCT/US99/17130

EARLIER FILING DATE: 1999-07-29

EARLIER APPLICATION NUMBER: 60/094, 657

EARLIER FILING DATE: 1998-07-30

EARLIER APPLICATION NUMBER: 60/095, 486

EARLIER FILING DATE: 1998-08-05

EARLIER APPLICATION NUMBER: 60/096, 319

EARLIER FILING DATE: 1998-08-12

EARLIER APPLICATION NUMBER: 60/095, 454

EARLIER FILING DATE: 1998-08-06

EARLIER APPLICATION NUMBER: 60/095, 455

EARLIER FILING DATE: 1998-08-06

NUMBER OF SEQ ID NOS: 376

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 132

LENGTH: 709

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (189)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

LOCATION: (414)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

LOCATION: (31)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

LOCATION: (438)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

LOCATION: (641)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

```
; LOCATION: (643)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (696)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-132

Query Match      82.8%; Score 3625.5; DB 2; Length 709;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAGINCGVSIALGLVLLGAAARLPRGAFAFIALPRESNITVLIKLGPTLLAKPCYIV 60
DB 1 MAGINCGVSIALGLVLLGAAARLPRGAFAFIALPRESNITVLIKLGPTLLAKPCYIV 60
QY 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLOPSTSLPT 120
DB 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLOPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQBEVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGBSATTMSANY 240
DB 181 GTVSRIRKQBEVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGBSATTMSANY 240
QY 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
QY 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSKFVPGCFVCLSERTCSNLTITSGSKHKISFLCDDLTRLMMANVEKTSCTDH 420
DB 361 PRPVKQSKFVPGCFVCLSERTCSNLTITSGSKHKISFLCDDLTRLMMANVEKTSCTDH 420
QY 421 RYCCOKKSYSLQVPSDILHLPELHDFESWKLVPKRLSLVLPACKLOQHTHEKPCNTSF 480
DB 421 RYCCOKKSYSLQVPSDILHLPELHDFESWKLVPKRLSLVLPACKLOQHTHEKPCNTSF 480
QY 481 SYLVASAIPOSQDLYFGSFCPGGSIKQIQVKONISVTLRTFAPSFXQEASROGLTVSFLPY 540
DB 481 SYLVASAIPOSQDLYFGSFCPGGSIKQIQVKONISVTLRTFAPSFXQEASROGLTVSFLPY 540
QY 541 FKKEGVFTVPTDYSKYVLTLPNMDRGPLSLTSVSMNISVPRDQVACLTFFKERSGVVQC 600
DB 541 FKKEGVFTVPTDYSKYVLTLPNMDRGPLSLTSVSMNISVPRDQVACLTFFKERSGVVQC 600
QY 601 TGRAFMIIOEORTRAEBIFSLDEDEVLPKPSFHHHSFWNINSQCTSGKOLDILFSVLT 660
DB 601 TGRAFMIIOEORTRAEBIFSLDEDEVLPKPSFHHHSFWNINSQCTSGKOLDILFSVLT 660
QY 661 PRYVDLTVILIAAVGGVLLSALGLLICVAKKKKKTKNGPAAVGIYNNX 710
DB 661 PRYVDLTVILIAAVGGVLLSALGLLICVAKKKKKTKNGPAAVGIYNNX 710

RESULT 2
US-09-489-847-273
; Sequence 273, Application us/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
```

```
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-273

Query Match      49.2%; Score 2155; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 5,4e-205;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCGVSIALGLVLLGAAARLPRGAFAFIALPRESNITVLIKLGPTLLAKPCYIV 60
DB 1 MAGINCGVSIALGLVLLGAAARLPRGAFAFIALPRESNITVLIKLGPTLLAKPCYIV 60
QY 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLOPSTSLPT 120
DB 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLOPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQBEVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGBSATTMSANY 240
DB 181 GTVSRIRKQBEVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGBSATTMSANY 240
QY 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
QY 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLASVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEYFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSKFVPGCFVCLSERTCSNLTITSGSKHKISFLCDDLTRLMMANVEK 413
DB 361 PRPVKQSKFVPGCFVCLSERTCSNLTITSGSKHKISFLCDDLTRLMMANVEK 413

RESULT 3
US-09-489-847-271
; Sequence 271, Application us/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
```

```

; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 271
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-271

Query Match      49.2%; Score 2155; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 6,1e-205;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKLGTPTLAKPCYVI 60
DB 30 MAGNCGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKLGTPTLAKPCYVI 89

QY 61 SKRITMLSTKSGRIYVTFSCSPENHFVIEIQKIDCMGSGPCPGEVOLQPSSTLPT 120
DB 90 SKRITMLSTKSGRIYVTFSCSPENHFVIEIQKIDCMGSGPCPGEVOLQPSSTLPT 149

QY 121 LNRFTMDVKAKHSIGLEQSIPLRLQIGSGCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 150 LNRFTMDVKAKHSIGLEQSIPLRLQIGSGCPDGVTHSISGRIDATVVRIGTFCSN 209

QY 181 GTVSRIKMQEGVKALHLPMFPHRNVSGFSIANRSSIRLCTIESEFEGSGATLMSANY 240
DB 210 GTVSRIKMQEGVKALHLPMFPHRNVSGFSIANRSSIRLCTIESEFEGSGATLMSANY 269

QY 241 PEGPEDELMTWQVVPALHRAVSFLNPNLSNCRKEERYEYIYPGSTNPEVFKLEDK 300
DB 270 PEGPEDELMTWQVVPALHRAVSFLNPNLSNCRKEERYEYIYPGSTNPEVFKLEDK 329

QY 301 QPGMAGNFNLSTGCGDDQASPGILRLQFVLVQHPNENKITYVVDLSNERAMSLTIE 360
DB 330 QPGMAGNFNLSTGCGDDQASPGILRLQFVLVQHPNENKITYVVDLSNERAMSLTIE 389

QY 361 PRPVKQSRKFPVPGCFVLESRTCSNLTLSGSKHKSIFLDDDLTRLMMNVEK 413
DB 390 PRPVKQSRKFPVPGCFVLESRTCSNLTLSGSKHKSIFLDDDLTRLMMNVEK 442

RESULT 4
US-09-489-847-272
; Sequence 272, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 272
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

US-09-489-847-272

Query Match      18.9%; Score 827; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1e-73;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 FEIALPRESNITVLIKLGTPTLAKPCYIVSKRHTMLSTKSGRIYVTFSCSPENHF 89
DB 1 FEIALPRESNITVLIKLGTPTLAKPCYIVSKRHTMLSTKSGRIYVTFSCSPENHF 60

QY 90 VIEIQKIDCMGSGPCPGEVOLQPSSTLPTLNRFTMDVKAKHSIGLEQSIPLRLQI 149
DB 61 VIEIQKIDCMGSGPCPGEVOLQPSSTLPTLNRFTMDVKAKHSIGLEQSIPLRLQI 120

QY 150 GPGESECPDGVTHSISGRIDATVVRIGTFCSNGTVSRIM 188
DB 121 GPGESECPDGVTHSISGRIDATVVRIGTFCSNGTVSRIM 159

RESULT 5
US-09-341-461-2
; Sequence 2, Application US/09341461
; Patent No. 6586389
; GENERAL INFORMATION:
; APPLICANT: Hammond, Timothy G.
; APPLICANT: Verronst, Pierre J.
; TITLE OF INVENTION: Cubilin Protein, DNA Sequences Encoding Cubilin
; FILE REFERENCE: D6148
; CURRENT APPLICATION NUMBER: US/09/341,461
; CURRENT FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: PCT/US99/01259
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 40
; SEQ ID NO 2
; LENGTH: 3623
; TYPE: PRT
; ORGANISM: rat
; FEATURE:
; OTHER INFORMATION: amino acid sequence of rat cubilin protein
US-09-341-461-2

Query Match      3.2%; Score 141; DB 2; Length 3623;
Best Local Similarity 18.9%; Pred. No. 0.0035;
Matches 164; Conservative 85; Mismatches 315; Indels 306; Gaps 36;

QY 70 IKSGRIYVTFSCSPENHFVIEIQKIDCMGSGPCPGEVOLQPSSTLPTLNRFTI 126
DB 2306 VSSRERITLKFCTGCGSSYMFKAKYSIASCGTVS-GDSGVISIGVPTLPYANNVCQ 2364

QY 127 WDVAHKSIGLEQSIPLRLQIGSGCPDGVTHSISGRIDATVVRIGTFCSNGT 182
DB 2365 WPIRGLPGGYLTLSFEDFNL-QSSPG--CTMDPVEIWEANGTSGRV-----LGRYCGNST 2415

QY 183 VSRIKMQEGVKALHLPMFPHRNVSGFSIANRSSIRLCTIESEFEGSGATLMSANY 242
DB 2416 PSSVDTSSNV-ASVKFTDGSVTVASGRFLPKFS--ROYCGDDL--GGPTGTFTSPNVPN 2470

QY 243 GPPEDELMTWQVVPALHRAVSFLNPNLSN-----GNNAGNFNLSTGCGDDQASPGI 277
DB 2471 PNPGRICENTYIYQEGRRIVLTFTNLRISTQPSCSNBGLIVPNCISNSPFLQKLSRV 2530

QY 278 EERVEYIYPGSTNPEVFKLEDKQ-----GNMAGNFNLSTGCGDDQASPGI 325
DB 2531 NVTNEFKSGNTMK-VVFETDGSRPYGFTASYSTEDAVCGFLPSVSG--GNFSSPGY 2587

QY 326 -----LRLOQVVLVQHPNENKITYVVDLSNERAMSLTIEPRPVKQSRKFPVPGCFV 376
DB 2588 NGIRDYARNLDCEWTLSNPNRENSSISYFLFLEISTEGQDCT-----FD 2631

QY 377 CLESRTCSNLTLSGSKHKSIFLDDDLTRLMMNVEKTICTDHRVCORKSYSIQVPSDI 436
DB 2632 VLRFVVDADQGPL-----IEK-----FC-----SLSATPAP 2657
```

```
QY 437 LHP-----VELHDF-----SM 448
Db 2658 LVIPPOVWIRFVSNERVEYGFYEYSFTDGGIRTDNGVISPYNPLNSAATGCSW 2717
QY 449 KLVLPKD-----RLSLVLPVPAQKLOQHTHEKPCNTSFYLVAASAIPOSOLYGFSGPCGS 503
Db 2718 LLKAPGGTITLTLSDFLLEA-----GPTCTSDSVTVRANGDSPSPVIGRCQGSV 2766
QY 504 IKQIQVKN-ISVLTRTAPSEFOASROGLTVSTIPIFKEE-----GVFTVTPPTKSKV 557
Db 2769 PRPIQSGSNOLIVFNT-----NNQGTFRGYATWTNALCGCGTFCGSANGT----- 2815
QY 558 YLRPNMNRGLPSLTVSW-----NISVPRQVACLTFEKRSG----- 596
Db 2816 -IKSGWQTPENSRCSWTYITGDSKGEISFDSNFIIPSSDSCQNSFVWGRMLMI 2874
QY 597 -----VVCOTGRAPMIIOBORTABEIFSLDEVLPRKSFHHSPWVNIS 641
Db 2875 NKTLLATSCGVAPSPVITSGNIFPAVFOSEEMAQGFS-----SF---IS 2918
QY 642 NC-----SPTSGKQJDLFSVTLTPRTVDLTVILANVGGVLLISA----- 683
Db 2919 RCGRTFNTSPGDIISPFPKQYDNNMNCITYLIDADPOSTVILITFVSFGLBDRSAITGTCD 2978
QY 684 ---IGLI-----ICVKKKKKKTKNKGPV-VGIYNXNINTEMPPROKPKQGRK 727
Db 2979 GDGGLIIGRNLSTPLVTICGSETLRPLTVDPVLNPFSDATYTDPG----- 3027
QY 728 DNDSHVAVIEDTVYVGHLODSSGSPLOP 757
Db 3028 -----FKISYRAITCGGIYNSSGILRSP 3051
```

```
RESULT 6
US-09-902-540-11386
; Sequence 11386, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11386
; LENGTH: 590
; TYPE: PR
; ORGANISM: Myxococcus xanthus
US-09-902-540-11386
```

```
Query Match 2.9%; Score 126.5; DB 2; Length 590;
Best Local Similarity 20.3%; Pred. No. 0.0042;
Matches 99; Conservative 57; Mismatches 156; Indels 175; Gaps 23;
```

```
QY 4 LNCGVSIALLGVL-----LGAAR-----LPRGAAAF 30
Db 5 LNLGLPLALLGFLAQAADVDVWGRGTNRILNAARIVKANDYATGYWLLPFAAVTL 64
QY 31 EIALPRESNITVILKIGPTTLAKPCYIVISKRAITMISIKSGEIVTFSCQSEPN-- 88
Db 65 NLNVPARO-----FGLNSDLVLAHMGSRGNGIWTFFDTPLSRYNL 106
QY 89 -----FVIRIQKNID-----CNSGPCPG-EVOLOPSTSLPPTLKRTPIMDVK 130
Db 107 GDSITVTVRGRGLKAYAAQVRGSDPFCNSACPHNVEITLPLFGSSAKVYKTIIFDID 166
QY 131 AHKSIGLELQF-----SIPRLR-----QIGGESCPDGVTHSISGRI-----DATVVR 173
```

```
Db 167 FTEDVQYK-QVAVAGVVRPRLLSSMAVVTPESSRCPSGSPSLSDGVVLSSPAPTCGILVD 225
QY 174 ICTFCSNGTYSRIRIKQSGVKA---LHP--WFHPRNVSGFSIANRSIK---RLCTIES 225
Db 226 LMSYDASGVLPVPEVPEGORSARFTLRLPYWTGPTVYVYGASGVKRSKVGVRPCLVHV 285
QY 226 VFEBSGATLMSANVPESGFPEDELMTOFV-----VPAHLRASVSFLFNENLNCERKEBRV 281
Db 286 -----FP---FVRMYLIDSLAPAHL-----LNGAVIARYKEAES 318
QY 282 EYIIPGTTNBEVKLED---KQPGNAGFNISLOGCDODASPGILRLQFQ----- 331
Db 319 DVLLTG---KGEAVNLNEVLGAEOVRVTVGNSTGDI FGTAYTAKGPNAPFLRSEDEVKAGA 375
QY 332 -----VLVQHPONESKIKTVVDLSNRASLTIERPRPVQSKRFVPGC 374
Db 376 EQMLEGMEAVTANAGTLLVBDP-NDKGIVLVD-----EVGPASH-----BGL 418
QY 375 FVCLESR 381
Db 419 SELLPSSR 425
```

```
RESULT 7
US-08-276-967-2
; Sequence 2, Application US/08276967
; Patent No. 5851817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; TITLE OF INVENTION: Sperm
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,967
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UTSD:418\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2476 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-276-967-2
```

```
Query Match 2.7%; Score 118.5; DB 1; Length 2476;
Best Local Similarity 22.5%; Pred. No. 0.31;
Matches 78; Conservative 45; Mismatches 120; Indels 103; Gaps 18;
```

```
QY 530 RQGLTVSRTIPIFKKEGVTTVPDTSK-VYLRTPNWRGLPSLTVSVSNISVPRQVACL 588
```


Db 172 RQGNKNTFI---OPAGPFGISLNGEHYIFLETDFKFSQAGSFRVSRPFCAP--AVICV 226
Qy 589 TFEFERSGVQCOTGRAPFIIIOEORTRAEIR-----SLDEVDLPK-PSFHHSHWV 638
Db 227 TFTYHMYGL-----GQGTIKRLILGSPAGSPSSLMRWVGQSP-----WL 268
Qy 639 NISNCSPTSGKQ-LDLLFSVTLTPTVDLTVILIAAVGGVLLLSALGLIICVKKKKK 697
Db 269 NTSVTISGHQHPQWOLIFE-----ANRGTNAFVALGFLV-----304
Qy 698 TNKGPAVGIXNKNINTEMPPROPKFKQGRKNDSHVAV-IEDTMVYGHLLQDSSGFLO 756
Db 305 INHGTGSPSETSVSTKVPAPTE---KPTVPSIYIPIPEKDMVH-----MEK 350
Qy 757 PEVDYTPFOGTMCVCPSPPTICSRAPTAK-----LATEPP-----PRSPSES 803
Db 351 PIYHTEKP---TVPEKRTIIPTEKSTVPTKPTVFEKPTLPPEGPVPAERPTTPPEGA 407
Qy 804 EP-----YTFSHPNNGDVSSKDDIP---LXTOEPMEPAE 836
Db 408 VPPKGPVLTVMPTSHTEKSTVHTEKPLPTGKSTIPEKEMVPTK 453

RESULT 8
US-09-964-956-62
; Sequence 62, Application US/09964956
; Patent No. 6875570
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Miller, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grose, William M
; APPLICANT: Alsebrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Bugess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OR INVENTION: No. 6875570e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964, 956
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235, 631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235, 633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235, 808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237, 434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238, 321
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238, 399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238, 396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276, 667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294, 823

; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304, 868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 2766
; TYPE: PR
; ORGANISM: Rattus norvegicus
US-09-964-956-62

Query Match 2.7%; Score 118; DB 2; Length 2766;
Best Local Similarity 19.0%; Pred. No. 0.42;
Matches 140; Conservative 95; Mismatches 265; Indels 238; Gaps 34;

Qy 144 PRLQIGPASCPCDVTHSISGRIDATVIRIGTCNGTVSRIMQGVKXALHPWFHP 203
Db 1865 PKPPQCVPEKPPQALGSL--RTSADTSIRFTSPKLPBEGANRPFH-----1917
Qy 204 RNVGFSIANRSSIKRLCTIESVPEGSGATLMSANYPEGPEDEL--MTQFVVPALR 261
Db 1918 -----AVYLESDTSCPTTSRSPRGPEKGAHPANGSASPASR 1956
Qy 262 ASVGFNLNLSNCRKEERYEYVIGSTNPEVFKLEDKQGNMAGNFNLSLQG-CDODA 320
Db 1957 ASLALAGIRQ-----KQFTPG--RADLVSEATDP-----QGI CEKGA 1993
Qy 321 QSPGILRLQFQVLVQHPQNESNKIYVVDLSNERAMSLTIEBRPVQGRKFPVGGFV---376
Db 1994 EKK-----VSDPPQTNQLKIVEISERVPKACGDRPPESDRK---GGFLQNN 2040
Qy 377 CLESRTSSNLTLTSSGKHKISFLCDLTLNMMVECTISDTHRYQGRKYSIQVP---433
Db 2041 CQEKSAIRLRQSESSPEH--TFPPSQASQ---VEREI-----RMSFSMAKPATS 2086
Qy 434 -SDILHLPVEL-HDFSK-----LIVPKDRLSLVVPAOKLQOHTHEKPCNTSFSYLV 484
Db 2087 SSSSLQLPALKEPFGQKSGOMPASVGPKNQVIGL--AGESEPTTPRATYTS---2141
Qy 485 ASALPSQ-DVYFG-----SFCPGSIRKQIOYKONI SVTLRTFADF--XQEASRQGLTV 535
Db 2142 --MFAQFSSHFGREGSPSPSHSPQDPQ-----PAMGKXSEKTAKGV-2185
Qy 536 SFLPYKEBGFYTTPTKSKVYLRTPNMDRGLPSLTVSKW---NISVPDQVACLTF 591
Db 2186 -----NGQGVYSVKPLETSSKNL-SPVDRDVSADETSCLIPDKVYTRRQYCCBQSW 2238
Qy 592 KERSGVVCOTGRAPFIIIOEORTRAEIRFSLDE-----DVLKPSFH-----633
Db 2239 PHES-----TSFSSVKRIKSFENJANSRPTAKATBPPLSVSSKPPINRRSSGSIP 2291
Qy 634 -----HSFWNISNCSPTSGKQDLIFSVTLPRTVDLTVILIAAVGGVLLLSAL 684
Db 2292 SGSPSDMTSRSLRSJSSGSEBSGLLPQMTKSSSMTLV-----2335
Qy 685 GLIICVKKKKKTKNCPA-----VGIXNKNINTEMPPROPKFKQGRKNDSHVYA 735
Db 2336 -----SRQNPDPSTSNKSPSPDKSLVPVGIPTSTVS---PASPSSKNK-----2376
Qy 736 VIEDTMVYGHLLQDSSGFLOPEVDYTPFOGTMCVCPSPPTIC-----SRAPTA---K 787
Db 2377 -----SSVRAHAPSPVSSKLUOERRTLSPDLKLCNGEDDSASFGAVLPK 2422
Qy 788 LATEPPPRSPSESESEP 805
Db 2423 TQLEITPRRSKGSQATSP 2440

RESULT 9
US-08-755-587-189
; Sequence 189, Application US/08755587
; Patent No. 6045997
; GENERAL INFORMATION:

```

; APPLICANT: Futreal, Phillip A
; APPLICANT: Mooster, Richard F
; APPLICANT: Ashworth, Alan
; APPLICANT: Straton, Michael R
; TITLE OF INVENTION: Materials and methods relating to the
; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESS: Bell Selzer Park & Gibson
; STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,587
; FILING DATE: 25-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9523959.6
; FILING DATE: 23-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525555.0
; FILING DATE: 14-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9617961.9
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenneth D Sibley
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405-135
; INFORMATION FOR SEQ ID NO: 189:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-755-587-189

Query Match      2.6%; Score 114; DB 2; Length 1589;
Best Local Similarity 17.7%; Pred. No. 0.4;
Matches 134; Conservative 104; Mismatches 234; Indels 286; Gaps 34;

QY 198 LPWFHPNVSQF-----STANSSIKRLCIHSVEEGSATLMSANYPEGPEDEL 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 558 LPNINEMFGFCSALGTKLGSVNEALRKAMKLFSDIENSEPBAKVG--PRGSSS-- 612
QY 250 MTWQFVVAHLRASVPLNPNLSNCRKEER-----VEYIIPST-----TNPEVKL 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 613 -----AHHDSVAIVFKIKKQNTKESDEKSSKCQVTLQNNIEMTTCIFVGRNPEKYIK 664
QY 298 EDKQGNMAGNFNLSIQGCDODQSPGILRLQFVLVQHPQESNKIYV---VDLSNER 353
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 665 NTKHEDSTSSQORNNLHNSDSGMS-----TSGPVYIHKGDSDL----- 703
QY 354 AMSLTIEPRPVQSKRFVPGCFVCLSRTCSSNLTLTSGSKHKISFLCDDLRLMWNVEK 413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 704 -----PADQSGK-----CPESCTQYARENTQ-KENIS-----DLTCLIMKAE 742
QY 414 TISCTDHRVCQKRSYSLQVPSDILHLPEVLHDFSM----- 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 743 T-----CMKSSDKQLPSDKMEQNIKEFNISFQIASGKNTRVSKESLNSKSVNIFNRE 794
QY 449 -KLIVPKDRL-SLVLVPAOKLQOHT--HEKPCNTSFSYLVASATPSODLVGSPCPGGS 503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 795 TDELTVISDSLNSKILHGINDKAMTSCCHK-----ALSIKVPEDH--PIVT 840
QY 504 IKQIQVKONISVTLRTFAPSPFOEASROGLTVSF-----IPYKE--EGVFTVTPDTK 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 841 VSQLPAQGH-----PEYEIESTKEPTLLSFTTASGKVKYKINQESLIDKYNLFDFTQ 891

```

```

QY 555 SKVYLR--TPNWDRLGSLTSVSNMISVPRDOACLTFPEKERSGVVCQGRAPMIIQEQRT 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 892 ---YVRKTAFSPQSGSKPL-----KQSKKEJLAYER-----LEVTAS 925
QY 614 RAERI---FSLDEVLPRKPSFNHNSFWNINSCSPISGK-QLDILFVTLTPRTVDLTVI 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926 KCEMNQNVSKETEMLPQONHMYRQTEMLTSGTSKQVENIENNVEKNPR----- 978
QY 670 LIAAVGGVLLSALGLIICCVKKKKTKNGPVGIVYNNKINTEMPPQPKFQKGR-- 727
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 979 -----ICCIQSSYPVTDSALAVYTEDSRKTVRE--SLSKGRFWL 1019
QY 728 -----DNDSHVYVI-----EDTM 741
DB 1020 REQGDKLTNTITICEKEHTEPDAGNASHSLVITRETIDTNHVSNOVSTLSDPNV 1079
QY 742 VYGHLDPS-----SGSFLOPEVDYRPFQGTWGVCPSPBPPTICSRAPPAKATE 791
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1080 CHVLSQSSFGCHODMHNDSGYFLKNKIDS-----DVPPDM-----KNAEGNTI 1123
QY 792 EPPRSPPESESEPTTFSH-----PNNGDVSSKQTDI 823
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1124 SPRVSATKERNMLHPQTTINVCQKLETTNTPSHANKDVAI 1161

RESULT 10
US-09-538-092-999
; Sequence 999, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT FILING DATE: 2000-03-29
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurateSeqFormatter Version 0.9
; SEQ ID NO 999
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P23515
US-09-538-092-999

Query Match      2.6%; Score 112.5; DB 2; Length 440;
Best Local Similarity 19.1%; Pred. No. 0.062;
Matches 79; Conservative 62; Mismatches 144; Indels 129; Gaps 17;

QY 256 VPAHIRASVSFLNF-----NLNCRKEEVE---YIIPGTTNPEV-- 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 49 LPSGIQENIILHNLNYSNFTDLHNLQYTNLRTLDISNNLSESLPAHLPRSLNMMGSAAN 108
QY 295 --FKIEDQPGNMGNFNLISQGCDQ-----AOSPGILRLQFVLVQHPQES 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 NNITKLDDK--SDTAYQMLKLYDVKNNLEKRVLLIKNTLRSLLEVNLSSNKLMTVPTMP 166
QY 342 NKIYVVDLSNERAMSLTIEPRPVQSKRFVPGCFVCLSRTCSSNLTLTSGSKHKISFLC 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 167 SKLHIVDSNN--SLT-----QILPGLTL-----NLT----- 191
QY 402 DDLRLMNVKETSCTDHRVCQKRSYSLQVPSDILHLPEVLHDFSKWLVPKORLSLVL 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 192 -NLTHLYLHNNKFTFIPQSFQ--LFLQI--EITLYNNRWG--CDHKONITVLL 239

```

```

QY 462 VPAQLOOQHTHEKPCNTSSYLVA-SAISODLYPESFCBGSGIKOIQVKONISVLR-- 518
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KMMETKAAHVIGTPESTOISSLEKHHMYTPESGFTSSLTVSGMQVTVDITNISLVTPK 299
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 519 -----TEAPSFQXOASRQGLVSPfpy-----FKEE 544
      || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 VTKLPKQRYRKETTTGATLSKDTTFTSTDKAFPEPYEDTSTETINSHEAAATLTLLHLD 359
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 GVFT---VTPTDKSKVYLRTPNMMDRGLPELTSVSNWISVPRDQVACLPEKERS 595
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 GMVVTYTLSTSCSKSSPTPMTLSITSCMP-----VFSEMPQOSTLLNIMREET 407
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

: RESULT 11
: US-09-949-016-9282
: Sequence 9282, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CI:001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9282
: LENGTH: 451
: TYPE: PRT
: ORGANISM: Human
: US-09-949-016-9282

```

Query Match	2.6%	Score 112.5	DB 2	Length 451
Best Local Similarity	19.1%	Pred. No. 0.064		
Matches	79	Conservative	62	Mismatches 144; Indels 129; Gaps 17

QY	256	VPAHLRASVSFLNF-----	-NISNCRKEERVE--	YIIPGSTNPEV--	294
		::: ::: ::: :::	::: ::: ::: :::	::: ::: ::: :::	
DB	60	LPISGEOENIIHLNLSYNHFTDLHNLQTYTLRLRTLIDISNNRRESIPAHLPRLSMNMSAAN			119
QY	295	--FTLEDPQPGNMGNFNLISLGGCOD-----	AQSPGLRLQFOVYQHPQES		341
		::: ::: :::	::: ::: ::: :::		
DB	120	NNIKLLDCK--SDTAYQNMKLKLYDVSKNMLEKVVLIKNTLRSLLEVNLSSNKLMTVPITNMP			177
QY	342	NKIYVVDLSNERAMSLTIEPRPVKQSRKRVPGCFVLESRTCCSNMLTITSGSKHISFLC			401
		::: ::: ::: ::: ::: ::: :::	::: ::: ::: :::		
DB	178	SKLHIVDLSNN--SLT-----	QILPOTLI-----	NLT-----	202
QY	402	DDLRLMNMNVKETSCTDHRRCQKRSYSLOYVPSDLHLPEVLHDSWKLVPKQRLSLVL			461
		::: ::: ::: ::: ::: ::: :::			
DB	203	-NLHLIYLNHNKFTTIPQSPDQ--LFLQL-----	ETLYNNRWS--	CHKQNTYLL	250
		::: ::: ::: ::: ::: ::: :::	::: ::: ::: :::		
QY	462	VPAQKLOOHTHEKPCNTSFSYLVA-SALPSODLYGSEFCPGSGISQIOVKONISVTLR--			518
		::: ::: ::: ::: ::: ::: :::	::: ::: ::: :::		
DB	251	KMMMETKAHVIGTFCSTQISSLKEHNMYPTPSGFFSSLFTVSGMQTVDTINSLASVTPK			310
QY	519	-----TFAPSFQDEASQGLTVPFP-----			544
		::: ::: ::: ::: ::: ::: :::			
DB	311	VTKIPKOYRIKETTFGATLSKDTTFTSTDKAFVPPBEDTSTETINSHEAATAATLTIHIOD			370
		::: ::: ::: ::: ::: ::: :::			
QY	545	GVFT--VTDPDKSVYLRTENMWDGSLPSLISVSANISVPRDQVACTFFERS			595
		::: ::: ::: ::: ::: ::: :::	::: ::: ::: :::		
DB	371	GMVNTNLSLSTSKSPPTMTLSISGMN-----	NFSBMPQOSTLLNMREET		418

RESULT 12
DS-09-560-385A-12

```

; Sequence 12, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutland, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1694
; TYPE: prt
; ORGANISM: Rattus norvegicus
; US-09-560-385A-12

```

Query Match Similarity	2.6%;	Score 112.5;	DB 2;	Length 1694;
Best Local Similarity	18.8%;	Pred. No. 0.64;		
Matches	147;	Conservative	99;	Mismatches 251; Indels 285; Gaps 36;
QY	128	DVKAHKSIGLELOFSIPRLRQIGGESC	P	DVTHSIGRDLATVIRIGTFCSGNCTVRIK 187
DB	769	DLKGYTSLSTFLQ--RDLRENGCTED--	M	FMYVIGNKDASXDYIGMAVVDQLTVCY 822
QY	188	MOGEVKAHLHPWHPPNVS	S	FSLANSSIRKLIIESV-----PEG 229
DB	823	NLGDREAVQIDQVLTESQSEAVMDRK	F	FRITIQPAKLNTKEATSNKPAVAVDIEG 882
QY	230	EGSATLTMSAN-----YPEGPEDELT	M	WQVPAHLRAS-----VSFLN 268
DB	883	GSNTTLNLNDEDAVFVYGGVPPD-----	F	ELPSRLRFPFYKCIELDLNENVLSTYN 936
QY	269	---FNLSNCE-----RKEEREVYI	P	GSTTNPVEFKLEDKQGNAGNFNLSLQCGDQ 318
DB	937	FKTFENNTTEVEECCRRRKEESD	K	DKVQFEGTG---YARLPTQNAEPFNF----- 982
QY	319	DAQSPGLRLQFOV---LVQHPPONESKI	T	VYVDLSNERAM---SLTIEPRVRXQSRKVP 372
DB	983	-----IQTIQTTVDRLGLFFFAEKQDNFI-	S	LNTEGDNLVRYKLSNBPKEKGIKD-- 10322
QY	373	GCFLVCELSRITCSSNLTLSGSKHKIS	F	ISFLCDL--TRLMMNV-EKTIQCTDHRVCORKSYSL 430
DB	1033	-----TINDKDHSLITLTIGLQGRMM	I	NNNERSV----- 10622
QY	431	QVPSDILHLPELHDFESFKLI--VP---	K	ORLSLVVPAQKLQOQTHF----- 473
DB	1063	-----RIBEIRIDFSTVYLGIGIP	I	AIREFNFNSTPAFOGCMNKLKTSQVRLNDTVG 1115
QY	474	--KPCNTSFSGLVNSA-----IP	S	QDGLV---FGSFCPGSIKQIOYKON 512
DB	1116	VTKKCSQEDWK-LVNTASFSRGGQMS	F	TNLDVPSTDRQLSFGFTFPSPGTLNHQR-- 1172
QY	513	ISVTLRTFAPSFXQOASRQGLTVS	F	IFPYKBEQGVFTVPTDKSKVYLRTPNWDRGLPSLT 572
DB	1173	TSSLVLVLEDOHIELSTRD-----	S	NIPIFKSPGY-----MD 1205
QY	573	SVSNNISVPRDQVACLTFFKERSV	V	QCTGRAFNIIOEQTRABE---ITSLEBDVLPKP 629
DB	1206	GLLHHVS-----VISDTSGRLRL	I	LDQVLRNRQRLPSFSMAQOSARLG 1248
QY	630	SFHHSFWNVNISNCSPTSGKQLDL	F	SVTLTPTRVLD---TVLIIAIVGGVLLLSALGL 686
DB	1249	GGHEGCGISNV-----LVQCF	S	QSPSEVILDAKSYTKKDALGCG----- 1286
QY	687	IICCVKKKKKKTKNGPAVGI	LYNNK	INTEMPROPKKFOKGRKQNDSHVAVIEDTMVYGL 746
DB	1287	--CSL-----NKPPILMLF-----	K	SPERFKRGRIFN-----VNQL 1315
QY	747	LODSGSEFLQPEVDVTRP	FGCTMGV	CPSPPTICSR-----PTAKLATEEPPRPSRPE 800
DB	1316	MODAPQATNSTEA-----WODGR	S	CLPPLNTKASHRALQFGDSFTSHLLKLPOLLKPR 1370
QY	801	SE 802		

US-09-248-796A-20641
; Sequence 20641, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20641
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20641

Query Match 2.5%; Score 107.5; DB 2; Length 649;
Best Local Similarity 17.9%; Pred. No. 0.38;
Matches 102; Conservative 77; Mismatches 206; Indels 185; Gaps 22;

QY 348 DLS-----NERAMSLTIERPVKQS--RKVPFCFVCL-----ESRTCSNLTU 389
DB 94 DLSAYALDOQSCSEYENPTTQELIKIPGLVDLQKGLYMAEQEATSTTNLSL 153
QY 390 TSGSKHKSIFLCDDLTRLMNVEXTISCTDHRVC-----ORKSYSLOVPSDILHP 440
DB 154 YG-----ALIODDLNKLKDNNNISNGSNGRFALKSELETNGEQDSEKQTNNDKNOD 207
QY 441 VELHD-----FSWKLTVPKORLSLVPAQKLOQH----- 470
DB 208 IEMTDALNKEVQOEIIEFEQTLSPO---LKPAFLTCWHPPTTEVPAVGKODGTATI 263
QY 471 -----THEKPCNTSFYLVASAIKPSODLY-----FGSFCGGSIKQIQ 508
DB 264 NAIODGKIETRTLLPMLNLIKQINIVSWSPQSLIITCGANSELRAWSPDGKLG-- 320
QY 509 VKONISVTLRTFAPSPQASRQGLTVSPFYFKEGVFTVPDTKSKYVLRTPNMDRL 568
DB 321 ---NIASITVDEAISL--ENTKLISITISLSW--SPGKFLSIDSLNOYCI---WDGTT 370
QY 569 PSLTSVSWNISVPRDQVACLTFEKE--RSGVVCQTG--RAFMIIQORTAEIIFSLDBV 625
DB 371 ISLKIQIKNLEINDSVCSWLSSEDKFAVTNTNGIKIYDILQPPSSH---FGNQLDV 426
QY 626 LKPSFHHHSFVWNISNCSPTSQKQDLFSVTLTPRTYDVLVILAAVGGVLLLSALG 685
DB 427 QPIGLNGEHNISLMKLNPTV-----KLATCSDPDYSI---KVMSSSSSQECLD 474
QY 686 LIICCVKKKKKTKNGPAVGIV-----NXNI--NTEM 715
DB 475 LNIINEKXTYTLKHAAPMIGLIMLPDSNSNERSLLSVSMGILNIMDAKTSENIKSSBL 534
QY 716 PRQPKKQGRKNDSHYAVIEDTMVYGHLLQDSGSGFL----- 755
DB 535 FNNKDNFTNELKDDNNNDTIKQVLYFNAVL--SPDGKYLALGDDYCRVTIMVDVTHTYL 593
QY 756 -----QPEVDTYRPFQSTMGVC 772
DB 594 NEPKDLIRCKAVYKPEISAEDKLKATIGIC 623

Search completed: May 4, 2006, 22:27:24
Job time : 50 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:27:41 ; Search time 84 Seconds
(without alignments)
4158.397 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377
Sequence: 1 MAGNCGVSIALGVLLGA.....SSKOTDIFLXTQEPMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.dep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.dep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.dep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.dep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.dep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.dep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4375	99.9	836	5	US-10-781-564-1
2	4374	99.9	836	3	US-09-899-569A-4
3	3625.5	82.8	709	4	US-10-351-334-132
4	3454	78.9	749	3	US-09-899-569A-2
5	2155	49.2	414	4	US-10-351-334-273
6	2155	49.2	443	4	US-10-351-334-271
7	1779	40.6	343	4	US-10-028-072-162
8	1779	40.6	343	4	US-10-140-808-162
9	1779	40.6	343	4	US-10-121-049-162
10	1779	40.6	343	4	US-10-123-904-162
11	1779	40.6	343	4	US-10-140-470-162
12	1779	40.6	343	4	US-10-175-746-162
13	1779	40.6	343	4	US-10-176-918-162
14	1779	40.6	343	4	US-10-176-921-162
15	1779	40.6	343	4	US-10-137-865-162
16	1779	40.6	343	4	US-10-140-474-162
17	1779	40.6	343	4	US-10-142-431-162
18	1779	40.6	343	4	US-10-143-114-162
19	1779	40.6	343	4	US-10-142-419-162
20	1779	40.6	343	4	US-10-123-262-162
21	1779	40.6	343	4	US-10-142-423-162
22	1779	40.6	343	4	US-10-121-050-162
23	1779	40.6	343	4	US-10-141-755-162
24	1779	40.6	343	4	US-10-143-032-162
25	1779	40.6	343	4	US-10-123-108-162
26	1779	40.6	343	4	US-10-123-236-162
27	1779	40.6	343	4	US-10-123-261-162

28	1779	40.6	343	4	US-10-140-921-162	Sequence 162, App
29	1779	40.6	343	4	US-10-140-928-162	Sequence 162, App
30	1779	40.6	343	4	US-10-121-045-162	Sequence 162, App
31	1779	40.6	343	4	US-10-123-292-162	Sequence 162, App
32	1779	40.6	343	4	US-10-123-903-162	Sequence 162, App
33	1779	40.6	343	4	US-10-124-819-162	Sequence 162, App
34	1779	40.6	343	4	US-10-124-832-162	Sequence 162, App
35	1779	40.6	343	4	US-10-140-925-162	Sequence 162, App
36	1779	40.6	343	4	US-10-160-458-162	Sequence 162, App
37	1779	40.6	343	4	US-10-124-824-162	Sequence 162, App
38	1779	40.6	343	4	US-10-127-825A-162	Sequence 162, App
39	1779	40.6	343	4	US-10-127-829A-162	Sequence 162, App
40	1779	40.6	343	4	US-10-127-835A-162	Sequence 162, App
41	1779	40.6	343	4	US-10-127-839A-162	Sequence 162, App
42	1779	40.6	343	4	US-10-127-901A-162	Sequence 162, App
43	1779	40.6	343	4	US-10-128-633A-162	Sequence 162, App
44	1779	40.6	343	4	US-10-131-813A-162	Sequence 162, App
45	1779	40.6	343	4	US-10-131-818A-162	Sequence 162, App

ALIGNMENTS

RESULT 1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
US-10-781-564-1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; Sequence 1, Application US/10781564																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; Publication No. US20040247601A1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; GENERAL INFORMATION:																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; APPLICANT: QuiGley, James P.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; APPLICANT: Hooper, John D.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; APPLICANT: Testa, Jacqueline E.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; TITLE OF INVENTION: The Scripps Research Institute																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; FILE REFERENCE: 1361.036US1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; CURRENT APPLICATION NUMBER: US/10/781,564																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; PRIOR FILING DATE: 2004-02-18																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; PRIOR APPLICATION NUMBER: US 60/448,828																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; PRIOR FILING DATE: 2003-02-19																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; NUMBER OF SEQ ID NOS: 10																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; SOFTWARE: FastSeq for Windows Version 4.0																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; SEQ ID NO 1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; LENGTH: 836																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; TYPE: PRT																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; ORGANISM: Homo sapiens																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
US-10-781-564-1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Query Match																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Beet Local Similarity 99.9%; Score 4375; DB 5; Length 836;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
QY	1	MAGNCGVSIALGVLLGAARLP	GA	FE	IA	LP	RES	NI	TV	L	K	G	T	T	L	A	K	P	C	Y	I	V	I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							

QY 361 PRPVQSRKFVPGCFVCEESRTCSNLTLTSGSKHKISFLCDLRLMNMNVEKTIISCTDH 420
DB 361 PRPVQSRKFVPGCFVCEESRTCSNLTLTSGSKHKISFLCDLRLMNMNVEKTIISCTDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQKLOQHTHEKPCNTSF 480
DB 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQKLOQHTHEKPCNTSF 480
QY 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQYKONI SVTLRTFAPSFKOASROGLTJSFTPY 540
DB 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQYKONI SVTLRTFAPSFKOASROGLTJSFTPY 540
QY 541 FKEEGVFTVPTDKSKVYLRTFPMWDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQC 600
DB 541 FKEEGVFTVPTDKSKVYLRTFPMWDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQC 600
QY 601 TGRAFMIIOEORTRAEEIFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGKQDLLEFSVLT 660
DB 601 TGRAFMIIOEORTRAEEIFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGKQDLLEFSVLT 660
QY 661 PRTVDLTVILIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROPK 720
DB 661 PRTVDLTVILIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROPK 720
QY 721 KFOGRKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
DB 721 KFOGRKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836

RESULT 2
US-09-899-569A-4
Sequence 4, Application US/09899569A
Patent No. US20020142003A1
GENERAL INFORMATION:
APPLICANT: No. US20020142003Albert Schweizer
APPLICANT: Marwa Schertl-Mostageer
APPLICANT: Wolfgang Sommergruber
APPLICANT: Roger Absenher
TITLE OF INVENTION: Tumorsassoziiertes Antigen (B345)
FILE REFERENCE: 0652.2280001
CURRENT APPLICATION NUMBER: US/09/899,569A
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: DE 100 33 080.0
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: DE 101 19 294.0
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 60/243,158
PRIOR FILING DATE: 2000-10-25
PRIOR APPLICATION NUMBER: US 60/297,747
PRIOR FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 836
TYPE: PRT
ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 99.9%; Score 4374; DB 3; Length 836;
Best local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAGINGVSIATLIGVLLGAARLPKGAFAFEIALPRESNITVYIKGTPTLAKPCYIYI 60
DB 1 MAGINGVSIATLIGVLLGAARLPKGAFAFEIALPRESNITVYIKGTPTLAKPCYIYI 60
QY 61 SKRHITMLSIKSGERIVFTFSQSPENHVFIEIQKNIDCMGSCPGEVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSQSPENHVFIEIQKNIDCMGSCPGEVQLOPSTSLPT 120

DB 61 SKRHITMLSIKSGERIVFTFSQSPENHVFIEIQKNIDCMGSCPGEVQLOPSTSLPT 120
QY 121 LNRTITWVKAHKSIGLELOFSIPRLROIIGCEGSPDVTSHISGRIDATVIRICTPCSN 180
DB 121 LNRTITWVKAHKSIGLELOFSIPRLROIIGCEGSPDVTSHISGRIDATVIRICTPCSN 180
QY 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSGIRLCIIEVFEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSGIRLCIIEVFEGEGSATLMSANY 240
QY 241 PEGFEDELMTWQFVVPAPHLRASVFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
DB 241 PEGFEDELMTWQFVVPAPHLRASVFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
QY 301 QPNMAGNPNLSLOCCDDAOSPGILRLQFOLVQHPONESKITYVDLSNRAASLTIE 360
DB 301 QPNMAGNPNLSLOCCDDAOSPGILRLQFOLVQHPONESKITYVDLSNRAASLTIE 360
QY 361 PRPVQSRKFVPGCFVCEESRTCSNLTLTSGSKHKISFLCDLRLMNMNVEKTIISCTDH 420
DB 361 PRPVQSRKFVPGCFVCEESRTCSNLTLTSGSKHKISFLCDLRLMNMNVEKTIISCTDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQKLOQHTHEKPCNTSF 480
DB 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQKLOQHTHEKPCNTSF 480
QY 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQYKONI SVTLRTFAPSFKOASROGLTJSFTPY 540
DB 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQYKONI SVTLRTFAPSFKOASROGLTJSFTPY 540
QY 541 FKEEGVFTVPTDKSKVYLRTFPMWDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQC 600
DB 541 FKEEGVFTVPTDKSKVYLRTFPMWDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQC 600
QY 601 TGRAFMIIOEORTRAEEIFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGKQDLLEFSVLT 660
DB 601 TGRAFMIIOEORTRAEEIFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGKQDLLEFSVLT 660
QY 661 PRTVDLTVILIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROPK 720
DB 661 PRTVDLTVILIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROPK 720
QY 721 KFOGRKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
DB 721 KFOGRKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836

RESULT 3
US-10-351-334-132
Sequence 132, Application US/10351334
Publication No. US20040034196A1
GENERAL INFORMATION:
APPLICANT: Komatsoulis et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P2
CURRENT APPLICATION NUMBER: US/10/351,334
CURRENT FILING DATE: 2003-01-27
PRIOR APPLICATION NUMBER: 60/350,898
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/489,847
PRIOR FILING DATE: 2000-01-24
PRIOR APPLICATION NUMBER: PCT/US99/17130
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: 60/094,657
PRIOR FILING DATE: 1998-07-30
PRIOR APPLICATION NUMBER: 60/095,486
PRIOR FILING DATE: 1998-08-05
PRIOR APPLICATION NUMBER: 60/096,319
PRIOR FILING DATE: 1998-08-12

PRIOR APPLICATION NUMBER: 60/095,454
PRIOR FILING DATE: 1998-08-06
PRIOR APPLICATION NUMBER: 60/095,455
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 132
LENGTH: 709
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (189)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (275)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (414)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (438)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (641)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (643)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (696)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (697)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-351-334-132
Query Match 82.8%; Score 3625.5; DB 4; Length 709;
Best Local Similarity 98.3%; Pred. No. 5.5e-315;
Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

DB 361 PRPVQSKKFPVPGCFVCLSESTCSNLTLTSGSKHKISFLCDDLTRLMNVKTIISCTDH 420
QY 421 RYQKRSYSLOVPSDILHLPEVLHDFSWKLLVPRXDRSLVLPVQKLOOHTHEKPCNTSF 480
DB 421 RYQKRSYSLOVPSDILHLPEVLHDFSWKLLVPRXDRSLVLPVQKLOOHTHEKPCNTSF 480
QY 481 SYLVASAIPODLYFGSFCPEGSGIKQIOVKONISVTLRTFAPSPXQASRQGLTVSPFY 540
DB 481 SYLVASAIPODLYFGSFCPEGSGIKQIOVKONISVTLRTFAPSPXQASRQGLTVSPFY 540
QY 541 FKESGVFTVPTDTSKYVLRTPNMDRGLPSLTSVSMNLSVRDVAACLPFKESGVVQC 600
DB 541 FKESGVFTVPTDTSKYVLRTPNMDRGLPSLTSVSMNLSVRDVAACLPFKESGVVQC 600
QY 601 TGRAFMIIQEORTAEIEFSLDEBVLPKPSFHHHSFWNINSCSPTSQKQDLFLSVTLT 660
DB 601 TGRAFMIIQEORTAEIEFSLDEBVLPKPSFHHHSFWNINXSPTSQKQDLFLSVTLT 660
QY 661 PRTVDLTVILIAAVGGVLLISALGLITCCYKXKKKKTKNKPAGVINYXN 710
DB 661 PRTVDLTVILIAAVGGVLLISALGLITCCYKXKKKKXT-RGPAVGIVYGN 709

RESULT 4

US-09-899-569A-2
Sequence 2, Application US/09899569A
Patent No. US20020142003A1

GENERAL INFORMATION:

APPLICANT: No. US20020142003Albert Schweitzer

APPLICANT: Marwa Scherl-Mostagaeer

APPLICANT: Wolfgang Sommergruber

TITLE OF INVENTION: Tumorcassoziertes Antigen (B345)

FILE REFERENCE: 0652,2280001

CURRENT APPLICATION NUMBER: US/09/899,569A

PRIOR APPLICATION NUMBER: DE 100 33 080.0

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: DE 101 19 294.0

PRIOR FILING DATE: 2001-04-19

PRIOR APPLICATION NUMBER: US 60/243,158

PRIOR FILING DATE: 2000-10-25

PRIOR APPLICATION NUMBER: US 60/297,747

PRIOR FILING DATE: 2001-06-14

NUMBER OF SEQ ID NOS: 40

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 749

TYPE: PRT

ORGANISM: Homo sapiens

US-09-899-569A-2

Query Match 78.9%; Score 3454; DB 3; Length 749;
Best Local Similarity 90.0%; Pred. No. 1.4e-299;
Matches 674; Conservative 9; Mismatches 30; Indels 36; Gaps 6;

QY 67 MLSIKSGRIIVFTSCSPENHFVIEIQKINIDCSGCPFGEOVLOPSTSLPTNRTFI 126
DB 1 MLSIKSGRIIVFTSCSPENHFVIEIQKINIDCSGCPFGEOVLOPSTSLPTNRTFI 60
QY 127 WDVKAHKSIGLELOFSIPRLRQIGPGECPDGVTHSISGRIDATVVRIGTFCNSGVYRI 186
DB 127 WDVKAHKSIGLELOFSIPRLRQIGPGECPDGVTHSISGRIDATVVRIGTFCNSGVYRI 186
QY 187 KMOEGVMAHLHPFHPRNVSGFSIARSSIKRCLIESVEGEGSATLMSANYPEGPE 246
DB 187 KMOEGVMAHLHPFHPRNVSGFSIARSSIKRCLIESVEGEGSATLMSANYPEGPE 246
QY 247 DELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIGSTNPEVFKLEDKQPGNMA 306
DB 247 DELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIGSTNPEVFKLEDKQPGNMA 306
QY 301 QPQNMAGNFNLISQCCDDAOSPGILRLQFVLVQHPONESNKIIVVDSLNERAMSLTIE 360
DB 301 QPQNMAGNFNLISQCCDDAOSPGILRLQFVLVQHPONESNKIIVVDSLNERAMSLTIE 360
QY 361 PRPVQSKKFPVPGCFVCLSESTCSNLTLTSGSKHKISFLCDDLTRLMNVKTIISCTDH 420

```
QY 307 GNFNLSTLQCCDDAOSPGLRLQFOVLVOHPONESNKIYVVDLSNERAMSLTIEBRPVKQ 366
DB 241 GNFNLSTLQCCDDAOSPGLRLQFOVLVOHPONESNKIYVVDLSNERAMSLTIEBRPVKQ 300
QY 367 SRKPVGCFVCLSESTCCSNLTLTSGSKHKISFLCDDLRLMWNVEKTICTDHRVCQK 426
DB 301 SRKPVGCFVCLSESTCCSNLTLTSGSKHKISFLCDDLRLMWNVEKTICTDHRVCQK 360
QY 427 SYSLQVPEDILHLPELHDFSWKLLVPKDRSLVVPKQKQOHTHEKPCNTSFSLVAS 486
DB 361 SYSLQVPEDILHLPELHDFSWKLLVPKDRSLVVPKQKQOHTHEKPCNTSFSLVAS 420
QY 487 AIPSDILFGSFCPSGSIKQIOVKONI SVTLRTFAPSFQOZASROGLTUSFIPIYKEGCV 546
DB 421 AIPSDILFGSFCPSGSIKQIOVKONI SVTLRTFAPSFQOZASROGLTUSFIPIYKEGCV 480
QY 547 FTVTDTKSKYVLTLPNMDRGLPSLTYSWMNISVPRDQVACLTFPKERSGVVCGTGAFM 606
DB 481 FTVTDTKSKYVLTLPNMDRGLPSLTYSWMNISVPRDQVACLTFPKERSGVVCGTGAFM 540
QY 607 TIQERTAAEELFSLDEVDLPRKSFHHHSFWNINSCSPTSGKQDLFSVTLTTRTYDL 666
DB 541 TIQERTAAEELFSLDEVDLPRKSFHHHSFWNINSCSPTSGKQDLFSVTLTTRTYDL 600
QY 667 TVIILAAVGGVLLSALGLIICVYKKKKKTNGKPAVGIYXNINTEMPPROPKKFOKGR 726
DB 601 TVIILAAVGGVLLSALGLIICVYKKKKKTNGKPAVGIYXNINTEMPPROPKKFOKGR 659
QY 727 KDNDSHYVAVIDIYVYGHLLQDSSGSLQPEVDITRPFGTMYGC---PPSP----- 776
DB 660 -----ERTWT--PMCMQSSRT-----PMY--MGICYRIPAPSCSQRW 694
QY 777 PTICSRAPAKLATEEPSPSESESEP 805
DB 695 PTICSRAPAKLATEEPSPSESESEP 723
```

RESULT 5

```
US-10-351-334-273
; Sequence 273, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; PRIOR FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-351-334-273
```

Query Match 49.2%; Score 2155; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 1.2e-183;

```
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGNCGVSIALLGVLLGAARLPRGAFAEIALPRESNITVTLKGTPTLLAKPCYVI 60
DB 1 MAGNCGVSIALLGVLLGAARLPRGAFAEIALPRESNITVTLKGTPTLLAKPCYVI 60
QY 61 SKRHITMISISGERIVTFSCQSPENHFVIEIQKNDICMGSPCFGEVOLQPTSLPPT 120
DB 61 SKRHITMISISGERIVTFSCQSPENHFVIEIQKNDICMGSPCFGEVOLQPTSLPPT 120
QY 121 LNRFTIWDVKAHKISGLEFQPSIRPLROI GGESCPDGYTISISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKISGLEFQPSIRPLROI GGESCPDGYTISISGRIDATVVRIGTFCSN 180
QY 181 GTVSRITKQGEVKMALHLPWPHPRNVSGESIANSSIRLCLISVFEGESATLMSANY 240
DB 181 GTVSRITKQGEVKMALHLPWPHPRNVSGESIANSSIRLCLISVFEGESATLMSANY 240
QY 241 PEGFPEDELMTQEFVPAHLRASVSLNPNLSNCKEKEVEYYIPGSTNPEVFKLEDK 300
DB 241 PEGFPEDELMTQEFVPAHLRASVSLNPNLSNCKEKEVEYYIPGSTNPEVFKLEDK 300
QY 301 QPGNAGNFNLSTLQCCDDAOSPGLRLQFOVLVOHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNAGNFNLSTLQCCDDAOSPGLRLQFOVLVOHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSKRFVPGCFVCLSESTCCSNLTLTSGSKHKISFLCDDLRLMWNVEK 413
DB 361 PRPVQSKRFVPGCFVCLSESTCCSNLTLTSGSKHKISFLCDDLRLMWNVEK 413
```

RESULT 6

```
US-10-351-334-271
; Sequence 271, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; PRIOR FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 271
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-351-334-271
```

Query Match 49.2%; Score 2155; DB 4; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.3e-183;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGNCGVSIALLGVLLGAARLPRGAFAEIALPRESNITVTLKGTPTLLAKPCYVI 60
DB 30 MAGNCGVSIALLGVLLGAARLPRGAFAEIALPRESNITVTLKGTPTLLAKPCYVI 89
```

Qy	61	SRBHTMTLSIKSGERIVFTEFCOSPENNHEVIEIOXNIDCMSCGPCEVOLQOSTSLPT	120
Dp	90	SRBHTMTLSIKSGERIVFTEFCOSPENNHEVIEIOXNIDCMSCGPCEVOLQOSTSLPT	149
Qy	121	LNRTFIWDPKAKHSIGLELOFSIPRLROIGPGSCDPGYTHSIGRIDATVNRIGTFCSN	180
Dp	150	LNRTFIWDPKAKHSIGLELOFSIPRLROIGPGSCDPGYTHSIGRIDATVNRIGTFCSN	209
Qy	181	GTVSRIKMOEGVKMALHLPMFHPRNVSGSINRSSIKRLCITIESPFEBCGSAITLMSANY	240
Dp	210	GTVSRIKMOEGVKMALHLPMFHPRNVSGSINRSSIKRLCITIESPFEBCGSAITLMSANY	269
Qy	241	PEGPEDELMTQFVVPAPHLRASVSFLNLSNCKEERKEVEYIDGSTTNPVFKLEDK	300
Dp	270	PEGPEDELMTQFVVPAPHLRASVSFLNLSNCKEERKEVEYIDGSTTNPVFKLEDK	329
Qy	301	QPGNNAQFNLSLOCCDDOAOISPGILRLOFOVLVQHPQBNESNKITYVDLSNEPAMSITIE	360
Dp	330	QPGNNAQFNLSLOCCDDOAOISPGILRLOFOVLVQHPQBNESNKITYVDLSNEPAMSITIE	389
Qy	361	PRPVQSRKFEVGCFCVLESRTCSSNLTLTSGSKHISFLCDLTLRLMNVK	413
Dp	390	PRPVQSRKFEVGCFCVLESRTCSSNLTLTSGSKHISFLCDLTLRLMNVK	442

```

RESULT 7
US-10-028-072-162
Sequence 162, Application US/10028072
GENERAL INFORMATION:
  APPLICANT: Baker, Kevin P.
  APPLICANT: Beresini, Maureen
  APPLICANT: DeForge, Laura
  APPLICANT: Desnoyers, Luc
  APPLICANT: Filvaroff, Ellen
  APPLICANT: Gao, Wei-Qiang
  APPLICANT: Gerltsen, Mary E.
  APPLICANT: Goddard, Audrey
  APPLICANT: Gurduey, Austin L.
  APPLICANT: Sherwood, Steven
  APPLICANT: Smith, Victoria
  APPLICANT: Stewart, Timothy A.
  APPLICANT: Tunes, Daniel
  APPLICANT: Macnabene, Colin K
  APPLICANT: Wood, William
  APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24

```

[illegible]

PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150; Indels 0; Gaps 0;
Matches 341; Conservative 2; Mismatches 0;

QY 1 MAGLNCGVSIALLGVLLGAAFLPGAGAFRIALPRESNITVLLKGTPTLLAKPCYIVI 60
1 MAGLNCGVSIALLGVLLGAAFLPGAGAFRIALPRESNITVLLKGTPTLLAKPCYIVI 60
DB 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQNIDCMGSPCPFGFVQLQPTSLAPT 120
1 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQNIDCMGSPCPFGFVQLQPTSLAPT 120
QY 121 LNRFTIWDVKAKHSIGLELQPSIPRLQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
121 LNRFTIWDVKAKHSIGLELQPSIPRLQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 181 GTVSRIRKQEGVKALHLPMFHPRVVSGFSTANRSSIRLCLIESVFEGGSATLMSANY 240
1 GTVSRIRKQEGVKALHLPMFHPRVVSGFSTANRSSIRLCLIESVFEGGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCKEERVEYYIPGSTNPEVFKLBDK 300
241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCKEERVEYYIPGSTNPEVFKLBDK 300
DB 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCKEERVEYYIPGSTNPEVFKLBDK 300
241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCKEERVEYYIPGSTNPEVFKLBDK 300
QY 301 QPMAGNPNLSLQCCDDAOSPGLRLQFOYLVQHPNBSNK 343
301 QPMAGNPNLSLQCCDDAOSPGLRLQFOYLVQHPNBSNK 343
DB 301 QPMAGNPNLSLQCCDDAOSPGLRLQFOYLVQHPNBSNK 343
301 QPMAGNPNLSLQCCDDAOSPGLRLQFOYLVQHPNBSNK 343

RESULT 8

US-10-140-808-162
Sequence 162, Application US/10140808
Publication No. US20030017563A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C182
CURRENT APPLICATION NUMBER: US/10/140,808
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-808-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAAARLPRGAFAFIALPRESNTIVLKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAAARLPRGAFAFIALPRESNTIVLKLGTPTLLAKPCYIV 60

QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

QY 121 LNRFTIMDVKAHKSIGLELOFSIPRLRQIGEGSCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELOFSIPRLRQIGEGSCPDGVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGEGSATLMSANY 240

QY 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCERKEERYEYIPGSTTNPVEFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCERKEERYEYIPGSTTNPVEFKLBDK 300

QY 301 QPGMAGNFNLSLOGCDDQDQSPGILRLQFOVLVQHQPONESK 343
DB 301 QPGMAGNFNLSLOGCDDQDQSPGILRLQFOVLVQHQPONESSE 343

RESULT 9
US-10-121-049-162
; Sequence 162; Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAAARLPRGAFAFIALPRESNTIVLKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAAARLPRGAFAFIALPRESNTIVLKLGTPTLLAKPCYIV 60

QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

QY 121 LNRFTIMDVKAHKSIGLELOFSIPRLRQIGEGSCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELOFSIPRLRQIGEGSCPDGVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGEGSATLMSANY 240

QY 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCERKEERYEYIPGSTTNPVEFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCERKEERYEYIPGSTTNPVEFKLBDK 300

QY 301 QPGMAGNFNLSLOGCDDQDQSPGILRLQFOVLVQHQPONESK 343
DB 301 QPGMAGNFNLSLOGCDDQDQSPGILRLQFOVLVQHQPONESSE 343

RESULT 10
US-10-123-904-162
; Sequence 162; Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAAARLPRGAFAFIALPRESNTIVLKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAAARLPRGAFAFIALPRESNTIVLKLGTPTLLAKPCYIV 60

QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

QY 121 LNRFTIMDVKAHKSIGLELOFSIPRLRQIGEGSCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELOFSIPRLRQIGEGSCPDGVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKQEGVKALHLPMFHPNRVSGFSIANRSSIRLCIIESVFEGEGSATLMSANY 240

```
Db 181 GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Db 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVOHPONESNK 343
Db 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVOHPONESSE 343

RESULT 11
US-10-140-470-162
; Sequence 162, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoige, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPGAEAFIALPRESNITVLKLGPTTLAKRCYIV 60
Db 1 MAGNCGVSIALGLVLLGAARLPGAEAFIALPRESNITVLKLGPTTLAKRCYIV 60
QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGVQLQPSLSLPT 120
Db 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGVQLQPSLSLPT 120
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
Db 181 GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Db 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVOHPONESNK 343
Db 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVOHPONESSE 343
```

```
RESULT 12
US-10-175-746-162
; Sequence 162, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoige, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPGAEAFIALPRESNITVLKLGPTTLAKRCYIV 60
Db 1 MAGNCGVSIALGLVLLGAARLPGAEAFIALPRESNITVLKLGPTTLAKRCYIV 60
QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGVQLQPSLSLPT 120
Db 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGVQLQPSLSLPT 120
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
Db 181 GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Db 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVOHPONESNK 343
Db 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVOHPONESSE 343

RESULT 13
US-10-176-918-162
; Sequence 162, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
```

APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKCYIVI 60
DB 1 MAGNCGVSIALGLVLLGAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKCYIVI 60
QY 61 SKRHITMISKGRIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSISLPT 120
DB 61 SKRHITMISKGRIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSISLPT 120
QY 121 LNRFTIMDVKAHKSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIISVFEGESATLMSANY 240
DB 181 GTVSRIKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIISVFEGESATLMSANY 240
QY 241 PEGFPEDELMTWQFVVAHLRAVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
DB 241 PEGFPEDELMTWQFVVAHLRAVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
QY 301 QPGNMAGNFNLSLQCCDDAOSPGLRLQFVLVQHFNQNESNK 343
DB 301 QPGNMAGNFNLSLQCCDDAOSPGLRLQFVLVQHFNQNESNK 343

RESULT 14
US-10-176-921-162
Sequence 162, Application US/10176921
Publication No. US20030027276A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT FILING DATE: US/10/176, 921
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKCYIVI 60
DB 1 MAGNCGVSIALGLVLLGAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKCYIVI 60
QY 61 SKRHITMISKGRIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSISLPT 120
DB 61 SKRHITMISKGRIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSISLPT 120
QY 121 LNRFTIMDVKAHKSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIISVFEGESATLMSANY 240
DB 181 GTVSRIKQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIISVFEGESATLMSANY 240
QY 241 PEGFPEDELMTWQFVVAHLRAVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
DB 241 PEGFPEDELMTWQFVVAHLRAVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
QY 301 QPGNMAGNFNLSLQCCDDAOSPGLRLQFVLVQHFNQNESNK 343
DB 301 QPGNMAGNFNLSLQCCDDAOSPGLRLQFVLVQHFNQNESNK 343

RESULT 15
US-10-137-865-162
Sequence 162, Application US/10137865
Publication No. US20030032155A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C154

```

; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO: 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-162

```

```

Query Match      40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MAGLNCGVSIALLGVLLGAARLPARGAEAFETALPRESNITVLIKGTPTLAKPCYIV 60
Db      1 MAGLNCGVSIALLGVLLGAARLPARGAEAFETALPRESNITVLIKGTPTLAKPCYIV 60
QY      61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPFGEVQLQSTSLPT 120
Db      61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPFGEVQLQSTSLPT 120
QY      121 INRTIWDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121 INRTIWDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY      181 GTVSRIKQOEGVKMALHLPMFHPRVNVSFSTANRSSIKRLCTIESVFESEGSATLMSANY 240
Db      181 GTVSRIKQOEGVKMALHLPMFHPRVNVSFSTANRSSIKRLCTIESVFESEGSATLMSANY 240
QY      241 PEGPPEDEBLMTWQFVVPAPHLRASVSFLNENLNCERKEKEREVEYYIPGSTTNPEVFKLEDK 300
Db      241 PEGPPEDEBLMTWQFVVPAPHLRASVSFLNENLNCERKEKEREVEYYIPGSTTNPEVFKLEDK 300
QY      301 QPGNNAGNPNLSLQGCDDQDQSPGILRLQFQVLVOHPQNESNK 343
Db      301 QPGNNAGNPNLSLQGCDDQDQSPGILRLQFQVLVOHPQNESSE 343

```

Search completed: May 4, 2006, 22:31:46
 Job time : 86 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:32:01 ; Search time 17 Seconds
(without alignments)
2276.120 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377
Sequence: 1 MAGINCGVSIALGLVLLGA.....SSKDTDFLXTQEMERPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*
1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
6: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
7: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
8: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
9: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
10: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
11: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
12: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4375	99.9	836	11	US-11-050-857-989 Sequence 989, App
2	4374	99.9	836	11	US-11-050-857-988 Sequence 988, App
3	4048	92.5	770	11	US-11-050-857-252 Sequence 252, App
4	3625.5	82.8	709	11	US-11-229-769-132 Sequence 132, App
5	3395	77.6	649	11	US-11-050-857-990 Sequence 990, App
6	2597	59.3	495	11	US-11-050-857-1135 Sequence 1135, App
7	2168	49.5	487	11	US-11-050-857-253 Sequence 253, App
8	2155	49.2	414	11	US-11-229-769-273 Sequence 273, App
9	2155	49.2	443	11	US-11-229-769-271 Sequence 271, App
10	1779	40.6	343	9	US-10-131-826A-162 Sequence 162, App
11	1779	40.6	343	9	US-10-137-115B-162 Sequence 162, App
12	1779	40.6	343	9	US-10-137-873A-162 Sequence 162, App
13	1779	40.6	343	9	US-10-152-370-162 Sequence 162, App
14	1779	40.6	343	11	US-11-050-857-987 Sequence 987, App
15	1779	40.6	343	11	US-11-290-153-162 Sequence 162, App
16	1210	27.6	231	11	US-11-050-857-254 Sequence 254, App
17	1132	25.9	238	11	US-11-050-857-255 Sequence 255, App
18	827	18.9	159	11	US-11-229-769-272 Sequence 272, App
19	391	8.9	146	11	US-11-050-857-1136 Sequence 1136, App
20	138	3.2	4051	8	US-10-501-834-7 Sequence 7, Appli
21	127	2.9	595	11	US-11-264-096-2258 Sequence 2258, Ap

22	118.5	2.7	4059	8	US-10-501-834-6	Sequence 6, Appli
23	118	2.7	2766	9	US-10-877-346-62	Sequence 62, Appl
24	117	2.7	3623	9	US-10-995-561-553	Sequence 593, App
25	113	2.6	869	9	US-10-453-372-50	Sequence 50, Appl
26	113	2.6	2612	9	US-10-453-372-38	Sequence 38, Appl
27	113	2.6	2669	9	US-10-453-372-36	Sequence 36, Appl
28	113	2.6	3095	11	US-11-235-732-4	Sequence 4, Appli
29	113	2.6	3104	9	US-10-453-372-34	Sequence 34, Appl
30	113	2.6	3104	9	US-10-453-372-62	Sequence 62, Appl
31	113	2.6	3104	9	US-10-453-372-64	Sequence 64, Appl
32	113	2.6	3130	9	US-10-453-372-42	Sequence 42, Appl
33	113	2.6	3483	9	US-10-453-372-40	Sequence 40, Appl
34	113	2.6	3546	9	US-10-453-372-32	Sequence 32, Appl
35	112.5	2.6	1725	9	US-10-766-317-8	Sequence 8, Appli
36	112.5	2.6	1725	9	US-10-784-004-457	Sequence 457, App
37	112.5	2.6	1725	9	US-10-784-004-960	Sequence 960, App
38	112	2.6	1045	11	US-11-096-568A-30402	Sequence 30402, A
39	108	2.5	699	11	US-11-150-887-43	Sequence 43, Appl
40	106.5	2.4	1125	11	US-11-024-959-160	Sequence 160, App
41	105	2.4	4590	8	US-10-505-928-569	Sequence 569, App
42	104	2.4	766	11	US-11-189-301-21	Sequence 21, Appl
43	104	2.4	2768	9	US-10-510-101-72	Sequence 72, Appl
44	101.5	2.3	693	11	US-11-189-301-20	Sequence 20, Appl
45	101.5	2.3	927	11	US-11-189-301-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1																																						
US-11-050-857-989																																						
; Sequence 989, Application US/11050857																																						
; Publication No. US20060040278A1																																						
; GENERAL INFORMATION:																																						
; APPLICANT: Compugen Ltd																																						
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS																																						
; FILE REFERENCE: 1847.1005																																						
; CURRENT APPLICATION NUMBER: US/11/050.857																																						
; CURRENT FILING DATE: 2005-01-27																																						
; NUMBER OF SEQ ID NOS: 1150																																						
; SEQ ID NO 989																																						
; LENGTH: 836																																						
; TYPE: PRT																																						
; ORGANISM: Homo sapiens																																						
US-11-050-857-989																																						
Query Match																																						
Best Local Similarity 99.6%; Score 4375; DB 11; Length 836;																																						
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;																																						
QY	1	MAGINCGVSIALGLVLLGAARLP	GA	AF	FI	AL	PR	ES	NI	TV	L	K	G	T	P	T	L	A	K	P	C	Y	I	60														
DB	1	MAGINCGVSIALGLVLLGAARLP	GA	AF	FI	AL	PR	ES	NI	TV	L	K	G	T	P	T	L	A	K	P	C	Y	I	60														
QY	61	SKRHITMISIGSERIVTF	FS	CQ	SP	EN	H	F	V	I	E	I	Q	N	I	D	M	G	S	G	P	F	G	V	Q	L	P	T	S	L	I	P	T	120				
DB	61	SKRHITMISIGSERIVTF	FS	CQ	SP	EN	H	F	V	I	E	I	Q	N	I	D	M	G	S	G	P	F	G	V	Q	L	P	T	S	L	I	P	T	120				
QY	121	LNRFITMVKAKHSIGLE	LO	FS	I	PR	LO	I	G	S	E	C	P	D	V	T	H	S	I	G	R	I	D	A	T	V	R	I	G	T	P	C	S	N	180			
DB	121	LNRFITMVKAKHSIGLE	LO	FS	I	PR	LO	I	G	S	E	C	P	D	V	T	H	S	I	G	R	I	D	A	T	V	R	I	G	T	P	C	S	N	180			
QY	181	GTGVRIRKQEGVKALH	LP	FW	PR	N	V	S	G	F	S	I	A	N	S	S	I	R	L	C	I	I	S	V	F	E	G	S	A	T	L	M	S	A	N	240		
DB	181	GTGVRIRKQEGVKALH	LP	FW	PR	N	V	S	G	F	S	I	A	N	S	S	I	R	L	C	I	I	S	V	F	E	G	S	A	T	L	M	S	A	N	240		
QY	241	PEGFPEDELMTWQFV	PA	HL	RA	S	V	S	FL	N	F	N	S	C	R	E	K	E	R	V	E	Y	I	P	G	S	T	N	P	E	V	F	K	L	E	D	K	300
DB	241	PEGFPEDELMTWQFV	PA	HL	RA	S	V	S	FL	N	F	N	S	C	R	E	K	E	R	V	E	Y	I	P	G	S	T	N	P	E	V	F	K	L	E	D	K	300
QY	301	ORGWAGNFNLSLGGC	OD	AS	P	G	I	L	R	Q	P	V	I	O	H	P	O	N	E	N	K	I	Y	V	V	D	L	S	N	E	R	M	S	L	T	I	E	360
DB	301	ORGWAGNFNLSLGGC	OD	AS	P	G	I	L	R	Q	P	V	I	O	H	P	O	N	E	N	K	I	Y	V	V	D	L	S	N	E	R	M	S	L	T	I	E	360

```
Db 301 QPQWAGNPNLSLQCCDDAQS PGLRLQFOVLVQH PONESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Db 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Qy 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Db 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Qy 481 SYLVASAIPODLYFGSFCPGGSIKOIQVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Db 481 SYLVASAIPODLYFGSFCPGGSIKOIQVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDYSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Db 541 FKEBGFVTPTDYSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Qy 601 TGRAFMIIOEORTRAEELFSLDEDLVLPKPSFHHHSFWVNISNCSPTSGKQDLFLSVTLT 660
Db 601 TGRAFMIIOEORTRAEELFSLDEDLVLPKPSFHHHSFWVNISNCSPTSGKQDLFLSVTLT 660
Qy 661 PRVDLTVILAAVGGVLLLSALGLIICVYKKKKKTKNGPAGIYXNINTEMPROPK 720
Db 661 PRVDLTVILAAVGGVLLLSALGLIICVYKKKKKTKNGPAGIYXNINTEMPROPK 720
Qy 721 KFOGRKNDSHVAVIEDTWYGHLDSSGSFLQPEVDYRPPQGTMGVCPSPPTIC 780
Db 721 KFOGRKNDSHVAVIEDTWYGHLDSSGSFLQPEVDYRPPQGTMGVCPSPPTIC 780
Qy 781 SRAPTAKLATEPPRSPSPSESEBPTTSHPNNGDVSSKDTDIPLLXTOEPMERAE 836
Db 781 SRAPTAKLATEPPRSPSPSESEBPTTSHPNNGDVSSKDTDIPLLXTOEPMERAE 836
```

RESULT 2

```
US-11-050-857-988
; Sequence 988, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 988
; LENGTH: 836
; TYPE: prt
; ORGANISM: Homo sapiens
US-11-050-857-988
```

```
Query Match 99.9%; Score 4374; DB 11; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 MAGNCGVSIALGVLILGAARLPRGAEEIALPRESNITVLIKIGPTLLAKPCYVI 60
Db 1 MAGNCGVSIALGVLILGAARLPRGAEEIALPRESNITVLIKIGPTLLAKPCYVI 60
Qy 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPPGEGVQLDPTSLPT 120
Db 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPPGEGVQLDPTSLPT 120
Qy 121 INRTIIMDKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCGN 180
Db 121 INRTIIMDKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCGN 180
Qy 181 GTVSRIKQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGBSATLMSANY 240
Db 181 GTVSRIKQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGBSATLMSANY 240
```

```
Qy 241 PEGPEDELTMTQFVVPVPAHLRASVSFLNPNLSNCRKEERYEYIIPGSTNPEVEFKEDK 300
Db 241 PEGPEDELTMTQFVVPVPAHLRASVSFLNPNLSNCRKEERYEYIIPGSTNPEVEFKEDK 300
Qy 301 QPQWAGNPNLSLQCCDDAQS PGLRLQFOVLVQH PONESNKIYVVDLSNERAMSLTIE 360
Db 301 QPQWAGNPNLSLQCCDDAQS PGLRLQFOVLVQH PONESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Db 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Qy 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Db 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Qy 481 SYLVASAIPODLYFGSFCPGGSIKOIQVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Db 481 SYLVASAIPODLYFGSFCPGGSIKOIQVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDYSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Db 541 FKEBGFVTPTDYSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Qy 601 TGRAFMIIOEORTRAEELFSLDEDLVLPKPSFHHHSFWVNISNCSPTSGKQDLFLSVTLT 660
Db 601 TGRAFMIIOEORTRAEELFSLDEDLVLPKPSFHHHSFWVNISNCSPTSGKQDLFLSVTLT 660
Qy 661 PRVDLTVILAAVGGVLLLSALGLIICVYKKKKKTKNGPAGIYXNINTEMPROPK 720
Db 661 PRVDLTVILAAVGGVLLLSALGLIICVYKKKKKTKNGPAGIYXNINTEMPROPK 720
Qy 721 KFOGRKNDSHVAVIEDTWYGHLDSSGSFLQPEVDYRPPQGTMGVCPSPPTIC 780
Db 721 KFOGRKNDSHVAVIEDTWYGHLDSSGSFLQPEVDYRPPQGTMGVCPSPPTIC 780
Qy 781 SRAPTAKLATEPPRSPSPSESEBPTTSHPNNGDVSSKDTDIPLLXTOEPMERAE 836
Db 781 SRAPTAKLATEPPRSPSPSESEBPTTSHPNNGDVSSKDTDIPLLXTOEPMERAE 836
```

RESULT 3

```
US-11-050-857-252
; Sequence 252, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 252
; LENGTH: 770
; TYPE: prt
; ORGANISM: Homo sapiens
US-11-050-857-252
```

```
Query Match 92.5%; Score 4048; DB 11; Length 770;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 767; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 67 MUISIGRIYVFTFSCQSPENHFVIEIQKNIDMSGPCPPGEGVQLDPTSLPTLNRTFI 126
Db 1 MUISIGRIYVFTFSCQSPENHFVIEIQKNIDMSGPCPPGEGVQLDPTSLPTLNRTFI 126
Qy 127 MDVKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCNSGVSKI 186
Db 127 MDVKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCNSGVSKI 186
Qy 187 KMQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGBSATLMSANYEGPE 246
Db 187 KMQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGBSATLMSANYEGPE 246
```

Db 121 KMGGVKKALHLPMPHPRNVSGFSIANKSSIKRLCTIESVEGEGSATLMSANYPEGPE 180
QY 247 DELMTWQFVPAHLRASVSLNPNLSNCERKEEREVEYIIPGSTNPEVFKLEDKQPGMA 306
Db 181 DELMTWQFVPAHLRASVSLNPNLSNCERKEEREVEYIIPGSTNPEVFKLEDKQPGMA 240
QY 307 GNFPLSLQGCCQDQDOSPGILRLQOVLYOHQONENKLYVVDLSNERAMSLLTIEPRVQ 366
Db 241 GNFPLSLQGCCQDQDOSPGILRLQOVLYOHQONENKLYVVDLSNERAMSLLTIEPRVQ 300
QY 367 SRKFPVCGFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDHRVCRK 426
Db 301 SRKFPVCGFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDHRVCRK 360
QY 427 SYSIQVPSDILHLPEVLDHDFSMKLLVPRDLISLVPAQKIQQHTHEKPCNTSFSYLVAS 486
Db 361 SYSIQVPSDILHLPEVLDHDFSMKLLVPRDLISLVPAQKIQQHTHEKPCNTSFSYLVAS 420
QY 487 AIPBODLYFGSFCGGSIKQIOVKONISVTLRTAPASXQASRQGLTVSFIPIYKKEGV 546
Db 421 AIPBODLYFGSFCGGSIKQIOVKONISVTLRTAPASXQASRQGLTVSFIPIYKKEGV 480
QY 547 FTVVPDTRKSKYLLRTPMNDRGLPSLTSYMANISVPRDQVACLTFPKERSGVVCOGRAPM 606
Db 481 FTVVPDTRKSKYLLRTPMNDRGLPSLTSYMANISVPRDQVACLTFPKERSGVVCOGRAPM 540
QY 607 IIOGORTRAEIEFSLDEVDLPKPSFHHHSFWVNISNCSPTSGKQDLILFVTLTPRTVDL 666
Db 541 IIOGORTRAEIEFSLDEVDLPKPSFHHHSFWVNISNCSPTSGKQDLILFVTLTPRTVDL 600
QY 667 TVILIAVGGVLLLSNLGLICCVKKKKKTKNGRANGIYNXNINTEMPQPKPKFOGR 726
Db 601 TVILIAVGGVLLLSNLGLICCVKKKKKTKNGRANGIYNXNINTEMPQPKPKFOGR 660
QY 727 KDNDSHYAVIEDTMVYGHLLQDSSGFLQPEVDTPYRPFQGTMGVCPSPPTISRAPTA 786
Db 661 KDNDSHYAVIEDTMVYGHLLQDSSGFLQPEVDTPYRPFQGTMGVCPSPPTISRAPTA 720
QY 787 KLAIEPPRPSPESESEPYTFSHPNNGDVSSKDTDIPLLXTOBMEBAE 836
Db 721 KLAIEPPRPSPESESEPYTFSHPNNGDVSSKDTDIPLLXTOBMEBAE 770

RESULT 4
US-11-229-769-132
; Sequence 132, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PIDI1
; CURRENT APPLICATION NUMBER: US/11/229, 769
; CURRENT FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 709

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (189)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (275)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (414)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (438)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (641)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (643)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (696)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (697)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-229-769-132

Query Match 82.8%; Score 3625.5; DB 11; Length 709;
Best Local Similarity 96.3%; Pred. No. 2,1e-308;
Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAGLNCGVSTALGLVLLGAAARLPARGAEPIALPRESNTITVLKGTPTLLAKPYIYI 60
Db 1 MAGLNCGVSTALGLVLLGAAARLPARGAEPIALPRESNTITVLKGTPTLLAKPYIYI 60
QY 61 SKRHITMLSIKGERIVFTSCQSPENHPIYIEIQKIDCMGSPCPFGEVQLQPSLSLPT 120
Db 61 SKRHITMLSIKGERIVFTSCQSPENHPIYIEIQKIDCMGSPCPFGEVQLQPSLSLPT 120
QY 121 LNRFTINDVKAHKSIGLEQSTIRLRQIGBESCPGVTHSIGRIDATVVRIGTFCSN 180
Db 121 LNRFTINDVKAHKSIGLEQSTIRLRQIGBESCPGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMGGVKKALHLPMPHPRNVSGFSIANKSSIKRLCTIESVEGEGSATLMSANY 240
Db 181 GTVSRIKMGGVKKALHLPMPHPRNVSGFSIANKSSIKRLCTIESVEGEGSATLMSANY 240
QY 241 PEGPPEDELMTWQFVPAHLRASVSLNPNLSNCERKEEREVEYIIPGSTNPEVFKLEDK 300
Db 241 PEGPPEDELMTWQFVPAHLRASVSLNPNLSNCERKEEREVEYIIPGSTNPEVFKLEDK 300
QY 301 QPQGMAGNFPLSLQGCCQDQDOSPGILRLQOVLYOHQONENKLYVVDLSNERAMSLLTIE 360
Db 301 QPQGMAGNFPLSLQGCCQDQDOSPGILRLQOVLYOHQONENKLYVVDLSNERAMSLLTIE 360
QY 361 PRPVKQSKRVPGCFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDH 420
Db 361 PRPVKQSKRVPGCFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDH 420
QY 421 RYCORKSYSIQVPSDILHLPEVLDHDFSMKLLVPRDLISLVPAQKIQQHTHEKPCNTSF 480
Db 421 RYCORKSYSIQVPSDILHLPEVLDHDFSMKLLVPRDLISLVPAQKIQQHTHEKPCNTSF 480
QY 481 SYLVASAIIPBODLYFGSFCGGSIKQIOVKONISVTLRTAPASXQASRQGLTVSFIPIY 540

Db 481 SYVASAIPSDLYGSCFCPGGSIKQIQVKONISVTLRTFAPSFRQEASRQGLTFSIFPY 540
Qy 541 FKESGVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFKERSGVVQ 600
Db 541 FKESGVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFKERSGVVQ 600
Qy 601 TGRAMMIIOEORTRAEEIFSLDEEDVLPKPSFHHHSFWNISCSPTSGKQDLRSVTLT 660
Db 601 TGRAMMIIOEORTRAEEIFSLDEEDVLPKPSFHHHSFWNISCSPTSGKQDLRSVTLT 660
Qy 661 PRVVDLTVILIAAVGGVLLLSALGLIICVKKKKKKTKNKGAIVGINXN 710
Db 661 PRVVDLTVILIAAVGGVLLLSALGLIICVKKKKKKTKNKGAIVGINXN 709
RESULT 5
US-11-050-857-990
; Sequence 990, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 990
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-990

Query Match 77.6%; Score 3395; DB 11; Length 649;
Best Local Similarity 99.2%; Pred. No. 2,6e-288;
Matches 644; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 188 MOEGYKMLHLPMFPRNVSGFSINRSSIKRLCTIESEFEGEGATLMSANYPGPFED 247
Db 1 MOEGYKMLHLPMFPRNVSGFSINRSSIKRLCTIESEFEGEGATLMSANYPGPFED 60
Qy 248 ELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYITPGSTTNPEVFLBDKQGNMAG 307
Db 61 ELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYITPGSTTNPEVFLBDKQGNMAG 120
Qy 308 NFNLSLQCCDDAOSPGILRLQFOYLVOPONESNKIYVVDLSNERAMSLTIEPRPVQS 367
Db 121 NFNLSLQCCDDAOSPGILRLQFOYLVOPONESNKIYVVDLSNERAMSLTIEPRPVQS 180
Qy 368 RKFVGCFCVLESRTCCSNLTLTSGSKHISFLCDLTRLMMNVKTTISCTDHRVCORKS 427
Db 181 RKFVGCFCVLESRTCCSNLTLTSGSKHISFLCDLTRLMMNVKTTISCTDHRVCORKS 240
Qy 428 YSLQVPSDILHLPELHDFSMKLLVPAKRLSLVLPAPQKLOQHTHEKPCNTSFSYLVA 487
Db 241 YSLQVPSDILHLPELHDFSMKLLVPAKRLSLVLPAPQKLOQHTHEKPCNTSFSYLVA 300
Qy 488 IPSQDLTYGSCFCPGGSIKQIQVKONISVTLRTFAPSFRQEASRQGLTFSIFYEKEGV 547
Db 301 IPSQDLTYGSCFCPGGSIKQIQVKONISVTLRTFAPSFRQEASRQGLTFSIFYEKEGV 360
Qy 548 TTPPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFKERSGVVQOTGRAMI 607
Db 361 TTPPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDQVACLTFKERSGVVQOTGRAMI 420
Qy 608 IOEORTRAEEIFSLDEEDVLPKPSFHHHSFWNISCSPTSGKQDLRSVTLTPTPTVLT 667
Db 421 IOEORTRAEEIFSLDEEDVLPKPSFHHHSFWNISCSPTSGKQDLRSVTLTPTPTVLT 480
Qy 668 VILIAAVGGVLLLSALGLIICVKKKKKKTKNKGAIVGINXNINTEMPRQPKFQKGR 727
Db 481 VILIAAVGGVLLLSALGLIICVKKKKKKTKNKGAIVGINXNINTEMPRQPKFQKGR 540

Qy 728 DNDSHVAVIEDTWYGHLLDSSGSFLQPEVDTRYRPEQTMGVCPPSPPTICSRAPTAK 787
Db 541 DNDSHVAVIEDTWYGHLLDSSGSFLQPEVDTRYRPEQTMGVCPPSPPTICSRAPTAK 600
Qy 788 LATEPPRSPPESESEBYTFSHPNNGDVSSKQTDIPLLXTOEPMEPAE 836
Db 601 LATEPPRSPPESESEBYTFSHPNNGDVSSKQTDIPLLXTOEPMEPAE 649
RESULT 6
US-11-050-857-1135
; Sequence 1135, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1135
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-11-050-857-1135

Query Match 59.3%; Score 2597; DB 11; Length 495;
Best Local Similarity 99.4%; Pred. No. 1e-218;
Matches 492; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 342 NKIYVVDLSNERAMSLTIEPRPVQSKRFVPGCFVLESRTCCSNLTLTSGSKHISFLC 401
Db 1 NKIYVVDLSNERAMSLTIEPRPVQSKRFVPGCFVLESRTCCSNLTLTSGSKHISFLC 60
Qy 402 DDLRLMMNVKTTISCTDHRVCORKSYSLOVPSDILHLPELHDFSMKLLVPAKRLSLV 461
Db 61 DDLRLMMNVKTTISCTDHRVCORKSYSLOVPSDILHLPELHDFSMKLLVPAKRLSLV 120
Qy 462 VPAKRLQOHTHEKPCNTSFSYLVAALPSQDLTYGSCFCPGGSIKQIQVKONISVTLRTFA 521
Db 121 VPAKRLQOHTHEKPCNTSFSYLVAALPSQDLTYGSCFCPGGSIKQIQVKONISVTLRTFA 180
Qy 522 PSFYQEASRQGLTFSIFYEKEGVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSWNISVP 581
Db 181 PSFYQEASRQGLTFSIFYEKEGVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSWNISVP 240
Qy 582 RDQVACLTFKERSGVVQOTGRAMIIOEORTRAEEIFSLDEEDVLPKPSFHHHSFWNIS 641
Db 241 RDQVACLTFKERSGVVQOTGRAMIIOEORTRAEEIFSLDEEDVLPKPSFHHHSFWNIS 300
Qy 642 NCSPTSGKQDLRSVTLTPTPTVLTIVILIAAVGGVLLLSALGLIICVKKKKKKTKNG 701
Db 301 NCSPTSGKQDLRSVTLTPTPTVLTIVILIAAVGGVLLLSALGLIICVKKKKKKTKNG 360
Qy 702 PAVGIVGINXNINTEMPRQPKFQKGRKNDSHVAVIETWYGHLLDSSGSFLQPEVD 761
Db 361 PAVGIVGINXNINTEMPRQPKFQKGRKNDSHVAVIETWYGHLLDSSGSFLQPEVD 420
Qy 762 YRPFQTMGVCPPSPPTICSRAPTAKLATEPPRSPPESESEBYTFSHPNNGDVSSKDT 821
Db 421 YRPFQTMGVCPPSPPTICSRAPTAKLATEPPRSPPESESEBYTFSHPNNGDVSSKDT 480
Qy 822 DIPLXTOEPMEPAE 836
Db 481 DIPLXTOEPMEPAE 495

RESULT 7
US-11-050-857-253

```
; Sequence 253, Application US/11050857
; Publication No. US20060040276A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 253
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-253

Query Match          49.5%; Score 2168; DB 11; Length 487;
Best Local Similarity 100.0%; Pred. No. 3,26-181;
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLVLLGGAARLPRGAFAFIALPRESNITVLIKLGTPTLAKPCYIV 60
DB 1 MAGNCGVSIALLVLLGGAARLPRGAFAFIALPRESNITVLIKLGTPTLAKPCYIV 60
QY 61 SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGPGCPGVOLQOSTSLPT 120
DB 61 SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGPGCPGVOLQOSTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGPEDELMTWQVVPVPAHLASVSLFNPNLNCERKEERYIIPGSTTNPVFKLEDK 300
DB 241 PEGPEDELMTWQVVPVPAHLASVSLFNPNLNCERKEERYIIPGSTTNPVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKSRKFVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMVNEKTIS 416
DB 361 PRPVKSRKFVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMVNEKTIS 416

RESULT 8
US-11-229-769-273
; Sequence 273, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PIDICI
; CURRENT APPLICATION NUMBER: US/11/229,769
; PRIOR FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
```

```
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-273

Query Match          49.2%; Score 2155; DB 11; Length 414;
Best Local Similarity 100.0%; Pred. No. 3,46-180;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLVLLGGAARLPRGAFAFIALPRESNITVLIKLGTPTLAKPCYIV 60
DB 1 MAGNCGVSIALLVLLGGAARLPRGAFAFIALPRESNITVLIKLGTPTLAKPCYIV 60
QY 61 SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGPGCPGVOLQOSTSLPT 120
DB 61 SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGPGCPGVOLQOSTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGPEDELMTWQVVPVPAHLASVSLFNPNLNCERKEERYIIPGSTTNPVFKLEDK 300
DB 241 PEGPEDELMTWQVVPVPAHLASVSLFNPNLNCERKEERYIIPGSTTNPVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKSRKFVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMVNEK 413
DB 361 PRPVKSRKFVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMVNEK 413

RESULT 9
US-11-229-769-271
; Sequence 271, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PIDICI
; CURRENT APPLICATION NUMBER: US/11/229,769
; PRIOR FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 271
; LENGTH: 443
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-271

Query Match      49.2%; Score 2155; DB 11; Length 443;
Best Local Similarity 100.0%; Pred. No. 3,8e-180;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGGAARLPFGAEAFETALPRESNITVLIKLGPTTLAKPCYIV 60
DB 30 MAGNCGVSIALGLVLLGGAARLPFGAEAFETALPRESNITVLIKLGPTTLAKPCYIV 89
QY 61 SKRHITMLSIKSGEIVTFPSCQSPENHFVIEIQNIDCMGSPCFGEVQLOPSTSLPT 120
DB 90 SKRHITMLSIKSGEIVTFPSCQSPENHFVIEIQNIDCMGSPCFGEVQLOPSTSLPT 149
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 150 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 209
QY 161 GTVSRIKMOEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESEVFBEGSATLMSANY 240
DB 210 GTVSRIKMOEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESEVFBEGSATLMSANY 269
QY 241 PEGFPEDELMTQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNEVFRLBDK 300
DB 270 PEGFPEDELMTQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNEVFRLBDK 329
QY 301 QPNMAGNPNLSLQCCDDAOSPGILRLQFOVLVOHPONESNKIYVVDLSNERAMSLTIE 360
DB 330 QPNMAGNPNLSLQCCDDAOSPGILRLQFOVLVOHPONESNKIYVVDLSNERAMSLTIE 389
QY 361 PRPVQSKRFVPGCVCLSESTCSNLTITSGSKIKISFLCDLTRLMANVEK 413
DB 390 PRPVQSKRFVPGCVCLSESTCSNLTITSGSKIKISFLCDLTRLMANVEK 442

RESULT 10
US-10-131-826A-162
; Sequence 162, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117

```

```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-162

Query Match      40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGGAARLPFGAEAFETALPRESNITVLIKLGPTTLAKPCYIV 60
DB 1 MAGNCGVSIALGLVLLGGAARLPFGAEAFETALPRESNITVLIKLGPTTLAKPCYIV 60
QY 61 SKRHITMLSIKSGEIVTFPSCQSPENHFVIEIQNIDCMGSPCFGEVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGEIVTFPSCQSPENHFVIEIQNIDCMGSPCFGEVQLOPSTSLPT 120
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 161 GTVSRIKMOEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESEVFBEGSATLMSANY 240
DB 161 GTVSRIKMOEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESEVFBEGSATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNEVFRLBDK 300
DB 241 PEGFPEDELMTQFVVPALRLASVSFLNPNLSNCRKERVEYYIPGSTTNEVFRLBDK 300
QY 301 QPNMAGNPNLSLQCCDDAOSPGILRLQFOVLVOHPONESNK 343
DB 301 QPNMAGNPNLSLQCCDDAOSPGILRLQFOVLVOHPONESSE 343

RESULT 11
US-10-973-115B-162
; Sequence 162, Application US/10973115B
; Publication No. US20060040351A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODIN
; FILE REFERENCE: 39870-3330R1C300C1

```

```

; CURRENT APPLICATION NUMBER: US/10/973,115B
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 10/145,747
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/028,072
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/581,742
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: PCT/US00/05746
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/135,736
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 60/123,090
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-973-115B-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1,9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAARLPFGAEAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLIGAARLPFGAEAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
QY 61 SKRHITMSTISGERIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSSTLLPT 120
DB 61 SKRHITMSTISGERIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSSTLLPT 120
QY 121 LNRFTFIMDVKAHKSIGLELOFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFIMDVKAHKSIGLELOFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQVVPVPAHLASVSFLNFNLSNCRKEERVEYIIPGSTTNPEVFKLEDK 300
DB 241 PEGFPEDELMTWQVVPVPAHLASVSFLNFNLSNCRKEERVEYIIPGSTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESNK 343
DB 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESSE 343

RESULT 12
US-10-137-873A-162
; Sequence 162; Application US/10137873A
; Publication No. US20060084138a1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Deenoysers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
```

```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C149
; CURRENT APPLICATION NUMBER: US/10/137,873A
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: 60/045911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-137-873A-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1,9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAARLPFGAEAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLIGAARLPFGAEAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
QY 61 SKRHITMSTISGERIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSSTLLPT 120
DB 61 SKRHITMSTISGERIVTFSCQSPENHFVIEIQKNIDMSGPCPFGEVOLQPSSTLLPT 120
QY 121 LNRFTFIMDVKAHKSIGLELOFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFIMDVKAHKSIGLELOFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQVVPVPAHLASVSFLNFNLSNCRKEERVEYIIPGSTTNPEVFKLEDK 300
DB 241 PEGFPEDELMTWQVVPVPAHLASVSFLNFNLSNCRKEERVEYIIPGSTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESNK 343
DB 301 QPGMAGNFNLSLOGCDODDQSPGILRLQFOVLVQHPONESSE 343

RESULT 13
US-10-152-370-162
; Sequence 162; Application US/10152370
; Publication No. US20060084139a1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Deenoysers, Luc
; APPLICANT: Filvaroff, Ellen
```



```

; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Maty E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P333ORIC407
; CURRENT APPLICATION NUMBER: US/10/152,370
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-152-370-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAARLPGAEAFETALPRESNITVYLKGTPTLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLIGAARLPGAEAFETALPRESNITVYLKGTPTLAKPCYIV 60
QY 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGCPFGVQLQPSLSLPT 120
DB 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGCPFGVQLQPSLSLPT 120
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYPGSTTNPDEVFKLBDK 300
DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYPGSTTNPDEVFKLBDK 300
QY 301 QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVQHPONESNK 343
DB 301 QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVQHPONESSE 343

RESULT 14
US-11-050-857-987
; Sequence 987, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; PRIOR FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 987
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-987

Query Match          40.6%; Score 1779; DB 11; Length 343;
```

```

Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAARLPGAEAFETALPRESNITVYLKGTPTLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLIGAARLPGAEAFETALPRESNITVYLKGTPTLAKPCYIV 60
QY 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGCPFGVQLQPSLSLPT 120
DB 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGCPFGVQLQPSLSLPT 120
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELOFSIPRLROIIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYPGSTTNPDEVFKLBDK 300
DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYPGSTTNPDEVFKLBDK 300
QY 301 QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVQHPONESNK 343
DB 301 QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVQHPONESSE 343

RESULT 15
US-11-290-153-162
; Sequence 162, Application US/11290153
; Publication No. US20060073568A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Betesini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flavioff, Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Maty E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P333ORIC321
; CURRENT APPLICATION NUMBER: US/11/290,153
; PRIOR FILING DATE: 2005-11-30
; PRIOR APPLICATION NUMBER: US/10/146,728
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
```


;; PRIOR APPLICATION NUMBER: 60/059352
;; PRIOR FILING DATE: 1997-09-19
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 162
;; LENGTH: 343
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-11-290-153-162

Query Match 40.6%; Score 1779; DB 11; Length 343;
Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKLGTPTLLAKPCYIVI 60
Db 1 MAGINCGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKLGTPTLLAKPCYIVI 60
QY 61 SKRHITMSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLPSTSLPT 120
Db 61 SKRHITMSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLPSTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLELOFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRTFIMDVKAHKSIGLELOFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLMSANY 240
Db 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTMOFVPAHLRASVSFLNPNLSNCRKEKREVEYYIPGSTNPEVFKLEDK 300
Db 241 PEGFPEDELMTMOFVPAHLRASVSFLNPNLSNCRKEKREVEYYIPGSTNPEVFKLEDK 300
QY 301 QPGNMAGNFNLSLQCCDDAOSPGILRLQFOVLVQHPONESNK 343
Db 301 QPGNMAGNFNLSLQCCDDAOSPGILRLQFOVLVQHPONESSE 343

Search completed: May 4, 2006, 22:32:43
Job time : 19 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model
Run on: May 4, 2006, 22:15:20 ; Search time 45 Seconds
(without alignments)
1787.494 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4377
Sequence: 1 MAGINCGVSTALGVLLGA.....SSKOTDIPLXTQEPMEAP 836

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	137.5	3.1	3623	2	T08618
2	121.5	2.8	986	2	T33135
3	118.5	2.7	2476	2	T34022
4	117.5	2.7	1464	2	S58984
5	117	2.7	2499	1	A30788
6	117	2.7	3623	2	T09456
7	113.5	2.6	1263	2	AH2011
8	113	2.6	1694	2	S50065
9	112.5	2.6	440	2	A39613
10	112	2.6	582	1	VCVDAR
11	112	2.6	1228	2	S46754
12	111.5	2.5	942	1	U01674
13	111	2.5	603	2	T24315
14	111	2.5	1185	2	T19212
15	111	2.5	3329	2	T42205
16	111	2.5	3329	2	T30904
17	111	2.5	4919	2	T31105
18	110.5	2.5	692	1	S59833
19	110.5	2.5	1027	2	T19173
20	110.5	2.5	1828	2	A40115
21	109.5	2.5	737	2	T31349
22	109	2.5	1224	2	T40765
23	108.5	2.5	5762	2	A41819
24	108	2.5	699	1	S44763
25	107.5	2.5	1721	1	I58902
26	106.5	2.4	2083	2	T42721
27	106	2.4	737	2	T16737
28	106	2.4	830	2	T37973
29	105.5	2.4	914	2	B96592

30	105.5	2.4	1272	2	C96637	hypothetical prote
31	105	2.4	1487	2	S15904	alpha-1 proteinase
32	104	2.4	1091	1	IUCHNL	neuronal cell adhesi
33	104	2.4	2180	2	A47651	zinc-finger protei
34	104	2.4	2464	1	ORMSP1	microtubule-associ
35	103.5	2.4	558	2	T37567	probable NADPH cyt
36	103.5	2.4	727	2	G01792	transcription fact
37	103.5	2.4	818	2	S57078	probable protein K
38	103.5	2.4	866	2	T01171	GI/S transition co
39	103.5	2.4	1163	2	I56126	lymphocyte function
40	103.5	2.4	4131	2	T21085	hypothetical prote
41	103	2.4	449	2	A55362	procollagen I C-pr
42	103	2.4	1565	2	A02135	polyketide synthas
43	103	2.4	2347	1	TVHURS	kinase-related pro
44	102.5	2.3	810	1	A33380	interleukin-4 rece
45	102	2.3	418	2	T29737	hypothetical prote

ALIGNMENTS

RESULT 1									
T08618									
intrinsic factor-B12 receptor CUBILIN precursor - rat									
C:Species: Rattus norvegicus (Norway rat)									
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004									
C:Accession: T08618									
R:Moestrup, S.K.; Kozyraki, R.; Kristiansen, M.; Kayser, J.H.; Rasmussen, H.H.; Brault									
J. Biol. Chem. 273, 5235-5242, 1998									
A>Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antidiode.									
A:Reference number: Z16459; MUID:98148073; PMID:9478979									
A:Accession: T08618									
A>Status: preliminary; translated from GB/EMBL/DBJ									
A:Molecule type: mRNA									
A:Residues: 1-3623 <MOE>									
A:Cross-references: UNIPROT:O70244; UNIPARC:UPI00000E776A; EMBL:AF022247; NID:93834379									
C:Genetics:									
A:Gene: CUBILIN									
C:Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology									
C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membran									
F:1-20/Domain: signal sequence #status predicted <SIG>									
F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>									
F:133-164/Domain: EGF homology <EGF1>									
F:436-467/Domain: EGF homology <EGF>									
Query Match									
Best Local Similarity 18.9%; Pred. No. 0.58;									
Matches 151; Conservative 77; Mismatches 283; Indels 289; Gaps 34;									
QY	70	IKSGERIVFTFS	CQSPENHFVIEIQKIDCN	SGPCPFGEVQLPST	--LLPTLNRT	126			
DB	2306	VSSRERIYTK	FHTDGGSYMFKAKYSI	ASCGTIVS	--GDSGVISIGVPTLVANNVFCQ	2364			
QY	127	MDVAAHSIG	LELFSTIRLQIGP	SGCPGCV	-----HISGGLDIAVVAIGFCSNGT	182			
DB	2365	WFTIGLGHVY	LTLSFEDPNTL	--QSSPG--CTKDVEIEMNHTSGV	-----LGRCGNST	2415			
QY	183	VSRIKMOGV	MAHLPMFHPRNVS	GFSIANRSSIKRLCTIE	SVFEGE--GSAITLMSAN	239			
DB	2416	PSVSDTSSNV	--ASVKFTLDSVTAS	GRFLQKSS--RQVC	-----GGLHPTGTFTSPN	2467			
QY	240	YPEGFPED	ELMTWQFVPAHLRASV	FLNPLNS	-----C	274			
DB	2468	YFNPNP	PHARICEMTITVOEG	RRIYLTFTNLRLSTQ	PSNSEHLLVFNGIRNSPILQKLC	2527			
QY	275	ERKKE	REVEYIRGSTNPEV	KLEDKP	-----GNMGNFLSLQGDQDAQS	322			
DB	2528	SRVAVT	VEFKSSGNTMK--VF	FTTGSRPYGGFT	RTASTEDAVCGGLPSPVSG--GNSS	2584			
QY	323	PGI	-----LRLQ	QVAVLPHQPN	SNKLIYVDLSNERAMS	LTIEPRPVKQSRKFPV	373		
DB	2585	PGVNGIRD	YVANNLDCENTL	SNPNRENSISIT	YFLDISHQDCT	-----	2629		

```

Qy 374 CFCVLESRTCSNNLTGSGSKHKSFLCDDLTRLMANVEKTIISCTDHRVCOKSKYSLOVP 433
Db 2630 -FDVLEFFVGVADGDL-----I-EK-----FC-----SLAP 2654
Qy 434 SDILHLP-----VELHDF----- 446
Db 2655 TAPLVIYPQWIIHFVSNERVEYTGFIIEVSFTDCGIRTDGNVSISSPNYPNLYSAMTH 2714
Qy 447 -SMKLVLVKDLSLVLAQAQLOQTHKEKPCNTSSYLYASAIPQODLYFGSCFGSGSIK 505
Db 2715 CSMWLKAPEGHITTLTFSDFLLEAH---PCTSTSVTVRNDSFGSPVIGRYCGQSVPR 2770
Qy 506 QIOVKON-ISVTLRTFASFQXOASRQGLTVSFIYFKE-----GVFTVTPDTSKYVL 559
Db 2771 PIQSSNOLDIYFNT-----NNQOTGRFVATWTNMGCGGTHSANGT-----I 2816
Qy 560 RTPNWDRLPLSLTYSW-----NISVPRDQVACL-TPEKERSG----- 596
Db 2817 KSPHWPQTFEPENSRCSMTVTITHESKHWEISFDSNFRIPSSDQCONSFVKVWEGRLMINK 2876
Qy 597 -----VVCQGRAFMIIOEORTRAEELFS----- 620
Db 2877 TLATSCGDVAPSPVLTSGNIFTAVFQSEBEMAAQGFSAFISRCGRTPNTSGDIISPNF 2936
Qy 621 -----LDED-----VLKPSFPH-----HHSFWMNISNCSPTSGKQ 650
Db 2937 PKQYDNMMNCTYLIDADQSLVILTFVSGFHLEDRSAITGTCHDGLHI-INKRNLSSTPL 2995
Qy 651 LDLEFVTLTPRTVDLTVIL 670
Db 2996 VTICGSETLRPLTVDGPIVL 3015

```

RESULT 2

```

T33135
hypothetical protein C45G7.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33135
R:Dante, M.; Wamsley, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C45G7.
A:Reference number: 221288
A:Accession: T33135
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-986 <DAN>
A:Cross-references: UNIPROT:O76355; UNIPARC:UPI000007ECAB; EMBL:AF067611; PIDN:AAC19183.
A:Experimental source: strain Bristol N2; clone C45G7
C:Genetics:
A:Gene: CESP:C45G7.5
A:Map position: 4
A:Introns: 373/3; 466/1; 524/3; 559/3; 722/2; 776/3; 839/2; 969/2
C:Superfamily: Caenorhabditis elegans hypothetical protein C45G7.5

```

Query Match 2.8%; Score 121.5; DB 2; Length 986;

Best Local Similarity 18.3%; Pred. No. 1.3; Mismatches 243; Indels 223; Gaps 32;

Matches 128; Conservative 101; Mismatches 243; Indels 223; Gaps 32;

```

Qy 127 MDVXAKSIGLELQFSIRP-----LRQIGPGECCPDG 158
Db 387 FETKMMVAVACELHMSVPREKRSQLECKLKEKLQALRENNKRLLYLRGMGHKT-PVK 445
Qy 159 VTHSGISGIDATVAVIGFCISNGYTSRIKMGQGVGKALHLPFPHRRNNSGFSIARRSSIK 218
Db 446 LAES-----STVENISKRETDTSVGRIQSYTRL--LPYLKETPEMHS----- 488
Qy 219 RLCTIESVFBGSGATLMSANY-----PEGFPED-----ELMTQFVVPALHLASVPLNF 269
Db 489 QYMTIQDL---EKSDVDESPPNPQKFLFDEPRAELLTLTANSGLVETSSSTLKI 545
Qy 270 NISNCCREERV-----EYVPGSTTNPVEFKLEDKQFGNNAFNISLQCCDDAOSP 323

```

```

Db 546 RVENLDNDEPFLPSALPIQVUPKNTSKPTAIGRLTARDADSPFIFYHLIPNCGTPRESSD 605
Qy 324 GI-LRLQGVVLVQHPO-NESNKIYVVDLSNERAMSLTIERPPVQSKRFVGGCFCLBSR 381
Db 606 NFNIIDAEFGELVEPEKVDPSKMTAEV-----CFIATQSK 641
Qy 382 TC-SSNLTLTGSGKHA---KISFLCDDLTRLMANVEKTIISCTDHRVCOKSKYSLOVPSDI 436
Db 642 NLDISEVFFDANSKNFKKVKTEFVG-----SDVSANSGKINNGFSNGNISVGDV 694
Qy 437 L---HLPEL-----HDPWKLVVKDRLSLVLPQO-----KLOQTHNE 473
Db 695 LDRVEIPNMLCAAGAOQDYEQ-----SLNFPAVYELGRDMSIPEGAVELKNNGE 746
Qy 474 KPCNTSFYLYASA---IPSDLY---FGSPCGSGSIQIOVKON-----ISVTLRTF 520
Db 747 -----LVANGKILDPQGVYTAELIGRSDGEGVQQLHINIRDKRLRYLSNRBF 797
Qy 521 APSEFOASRQGLTVSFIYFKEGVTVPDTSK---KYVLTTPMMDRLPSLTYSVS 575
Db 798 GANLEK-----FRQILEAIGKDDQAKQGLIHDPERKADKKNSTWTSV- 842
Qy 576 WNISVPRDQVACLTFEKRSGVVCOTGRAPMIIOEORTRAEELISLDEVLPKSFHHS 635
Db 843 -----CFYLTRONALIDENQASSIISPSNGHISKL-----HHI 875
Qy 636 FMV-NISNCSPT-----TSQKQDLDFSVTLTPRTVDLTVILAAVGGVLLSALG 685
Db 876 FKQVNDMCAFKRPAATSESTSSNSDIPLN-----TLIIQVVGILLIIMLALL 925
Qy 686 LIICCVK---KKKTKN---GPAVGIYXNINTEMPROP 719
Db 926 IYVCCVSRYQRYLKQKTDQLRCSASGYSYKSPNLIPEPPP 966

```

RESULT 3

```

T34022
zonadhesin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C:Accession: T34022
R:Hardy, D.M.; Garbers, D.L.
J. Biol. Chem. 270, 26025-26028, 1995
A:Title: A sperm membrane protein that binds in a species-specific manner to the egg ext
A:Reference number: 221464; MUID:96064658; PMID:7592795
A:Accession: T34022
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2476 <HAR>
A:Cross-references: UNIPROT:Q28983; UNIPARC:UPI000013C373; EMBL:U40024; NID:GL066465; PI
A:Experimental source: strain Meishan; testis
C:Genetics:
A:Gene: Zan
A:Function:
A:Description: may be involved in sperm adhesion to the zona pellucida

```

Query Match 2.7%; Score 118.5; DB 2; Length 2476;

Best Local Similarity 22.5%; Pred. No. 8.6; Mismatches 120; Indels 103; Gaps 18;

Matches 78; Conservative 45; Mismatches 120; Indels 103; Gaps 18;

```

Qy 530 RQGLTVSFIYFKEGVTVPDTSK-VYLRTPMWDRGLPSLTYSWNISVPRDQVACL 588
Db 172 RQGNKNTFI---QPAQPGISLNGGHYIFLETDFSQAGSFRVLSRPFCA-P-AVICV 226
Qy 589 TPEKERSGVVCOTGRAPMIIOEORTRAEELF-----SLDEVLK-PGFHHSRWV 638
Db 227 TETVHYMGL-----GQGTKLRLLSGSPASGSPSSLWERVGGQSP-----WL 268
Qy 639 NISNCSPTSGKO-LDLEFVTLTPRTVDLTVILAAVGGVLLSALGLIICCVKKKKK 697
Db 269 NTSVTIPSGHQDPMLQIE-----AVRGTTNAFVALGFLV----- 304
Qy 698 TNKGPVAVGIYXNINTEMPROPKFKQGRKNDNSHVAV-IEDTWVYGHLLDSSGSFLQ 756

```

```

Db      305 INHGTCRGPSETSVSTKQVAPTE-----KPTVSEIYITPTKPMWH-----MEX 350
Qy      757 PEVDYTRPPOGMGVCPPSPPTICSRAPPAK-----LATEEP-----PSPSESES 803
Db      351 PIYHTEKP---IVPTPEKPTIPTEKSTVPFKKPTVFKEPPLPEGGFTVPAERPTTPEGPA 407
Qy      804 EF-----YTFSPHNNGDVSSKOTDIP-----LXTOEPMEPAE 836
Db      408 VPKGPFTVLTWPTSTHTEKSTVTEKPIPLPTGKSTIPTEKPMVPTK 453

```

RESULT 4

```

558984
development protein tolkin (EC 3.4.24.-) - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C/Accession: S58984
R/Finelli, A.L.; Xie, T.; Bossele, C.A.; Blackman, R.K.; Padgett, R.W.
Genetics 141, 271-281, 1995
A/Title: The tolkin gene is a tolloid/BMP-1 homologue that is essential for Drosophila
A/Reference number: S58984; MUID:96042912; PMID:8536976
A/Accession: S58984
A/Molecule type: mRNA
A/Residues: 1-1464 <FIN>
A/Cross-references: UNIPROT:Q24132; UNIPARC:UPI00000820A1; EMBL:U34777; NID:g1002985; PI
A/Note: the authors did not translate the codon for residue 722
C/Genetics:
A/Gene: tolkin
A/Cross-references: FlyBase:FBgn0004885
C/Keywords: hydrolase; metalloproteinase; zinc
F/529-722/Domain: astacin homology <AST>
F/958-993/Domain: EGF homology <EGF>
F/1118-1153/Domain: EGF homology <EGF1>
F/614,618,624,673/binding site: zinc (His, His, His, Tyr) #status predicted
F/615/Active site: Glu #status predicted

```

```

Query Match      2.7%; Score 117.5; DB 2; Length 1464;
Best Local Similarity 19.3%; Pred. No. 4.6;
Matches 98; Conservative 63; Mismatches 161; Indels 185; Gaps 24;

Qy      105 PCGEVOLQSTSLPLTANRTFIMDVAKHSIGLELOFSIRLRQIGPSCGCDGVTHSIS 164
Db      1004 PNGIT-TSPSPFEMTPLEKCIWEIYAPPKRISLNTFTPD-----EGRAHQOS 1052
Qy      165 --GRIDATVV-----RIGTFCSNGTVSRIKQOEGVKALHLPWPHPRNV--SGFSI 211
Db      1053 DCGYDSVTVYYSKLGEMRLKRIGTFGSSSIPTATSESN---ALRIEFHSDKSIQSGFAA 1109
Qy      212 ANRSSIKR-----LCITISVF-----EGEGS-----ATL 235
Db      1110 VFTTIDECVANNGCCOHECRNTTIGYICMCHNGVSMHENGHDCKEGEKYEISAPFGTI 1169
Qy      236 MSANYPEGPEDELTMTQFV-VPAHLRASVSFLNPL-SNCERKEKRVETYPGATTNE 293
Db      1170 FSPNPDPDSYPPNADCVHMFITTPGH-RIKLIFNEPDVESHOCCTDNYAVY----- 1219
Qy      294 VFKELEKDPGNNAAGNFNLISLQCDQDAOSPGL-----RLOFOVLVQHPONESNKIYV 347
Db      1220 -----DSESSESVLGRFCGDKIPPI-----SSTSNQMWV 1251
Qy      348 DLSNERAMSLTEPRPVKQSRKFPVGCPCPLCSRCCSSNULTTSGSKHKISPL-----C 401
Db      1252 -----LKTDKNKKQNGFASHSTACGGYLRATISQVQOQFYSHARFGNDY 1295
Qy      402 DDLRLMMNVKTIISCTDHRVYQ-----RKYSY--LQVPSDILH----- 438
Db      1296 DD---GNDCEWTIAAPNSYVQLFLFTPDIESSERCFTDYQVSDIDVYGQYPMKG 1351
Qy      439 -----LVELHDFSMKLL-----VKDRLSLVLVPAQKLQDHT- EKPCNTSFSYL 483
Db      1352 QYCGNVLPQDINSMTSHLLVRFKTDGVSVMKGFASVYAVPVNPSGEYHSDHDEVDENSYSSE 1411

```

```

Qy      484 VASATPSQDLVYGFSGFCGGSIKQIOVK 510
Db      1412 MVTFP-----GSLKSYIE 1426

```

RESULT 5

```

A30788
mannose 6-phosphate receptor protein, cation-independent - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: A25908; A30788; S09404
R/Lobel, P.; Dahms, N.M.; Kornfeld, S.
J. Biol. Chem. 263, 2563-2570, 1988
A/Title: Cloning and sequence analysis of the cation-independent mannose 6-phosphate r
A/Reference number: A92706; MUID:88115411; PMID:2963004
A/Accession: A25908
A/Molecule type: mRNA
A/Residues: 1-2499 <LOB>
A/Cross-references: UNIPROT:P08169; UNIPARC:UPI000012F4B2; GB:J03527; NID:g162873; PIDN
R/Glickman, J.N.; Conibear, E.; Pearce, B.M.F.
EMBO J. 8, 1041-1047, 1989
A/Title: Specificity of binding of clathrin adaptors to signals on the mannose-6-phospi
A/Reference number: S09404; MUID:89305502; PMID:2545438
A/Contents: annotation: HA-II adaptor binding
C/Comment: This protein binds phosphorylated lysosomal enzymes and insulin-like growth
C/Superfamily: mannose 6-phosphate receptor, cation-independent; fibronectin type II r
C/Keywords: Golgi apparatus; membrane protein
F/1912-1951/Domain: fibronectin type II repeat homology <2F1>
F/2360-2363/Region: HA-II adaptor binding

```

```

Query Match      2.7%; Score 117; DB 1; Length 2499;
Best Local Similarity 19.8%; Pred. No. 11;
Matches 191; Conservative 103; Mismatches 343; Indels 326; Gaps 55;

Qy      1 MAGLNCGYSIALGLVLLGAARLPFGAAFEIALPRESNITVILKGTPTLL-----AKPC 56
Db      1710 MHGLACPRAGTAVCKRPVUGP-----PIDIRVAGPPLINIANEVINFESSITPC 1759
Qy      57 YIVSKRHITMLSIKSGRIVFTFSGQ-----SPE-----NHFVIEIKNIDC---M 100
Db      1760 ---LADRHFNVTSL-----ITFHCKRGVSGMTGPKLRTSCVDFPFMEPTLVCPDEVK 1809
Qy      101 SGCPFGEVQLOPSTSLPLTANRTFIMDVAKH-SIGLELOFSIRLRQIGGE-SCPDG 158
Db      1810 TDGCSLTBOQYYSNLSLSKSTPKVTRGPTYSVG-----CTAAAGIDEGCCKDG 1862
Qy      159 VTHSISGRIDATVVVIGTFCSNGTVSRIKQOEGVKALHLPWPHPRNVSGFSIANRSSIX 218
Db      1863 AVCLLSGSKGASFGRLA-----SKLDYRHODEAVIISYANGDTCP 1903
Qy      219 -----RLCITISVEEGEGSATLMSANYPEGPEDELTMTQFVPAHLRASVSFLNPLIS 272
Db      1904 PETEDGEPCCVPPFVNGK-----SYEECVESBARLM-----CATTANYORD 1945
Qy      273 N-----CEKEREVEYIIGSTTNPVEVFKLEDKQFGNMAAGNFNLISLQCDQDAOSGIIRL 328
Db      1946 HEWGFCKKSTSH-----RTSVIIFK-----CDEAD 1971
Qy      329 QFOVLVQHPONESNKIYVVDLSNERAMSLTEPR-----PVKQSRKFV----- 371
Db      1972 ---VGRPO-----VSEVRGCEVTTEMTKKVVCPRKKECKVQKRTYDRL 2017
Qy      372 ---PGCF-----VCLF-----SRTCSNLT-----TSGSKHKISF-----LC 401
Db      2018 SSLTGSWGFVHNGASYYINLCOKIYKGPQDCSERASVCKKSTSGEVQYLGVHTQKLVV 2077
Qy      402 DDLRLMMNVKTIISCTDHRVYQ-----RKYSY-LQVPSDILH-PVELHDPSSW 448
Db      2078 DD--RVIVTYSKGHYCGDNKTAASAVIELTCAKTVGRPSFTFPDVSCTYH-----FSW 2128
Qy      449 KLVVFKDRLSLVLVPAQKLQ---QHTHEKPC-TSFSYLVASATPSQDLVYGFSGFCGGSIK 505
Db      2129 ---DRAACAVKPGQEVQVQVNGTITTPAAGRFSFL-----GDITFKRFSASGDVR 2174

```

QY 506 -----OIOVKONISVTLRTAPAEFX-----QEAASROGLTVSPFIPEKEEG- 545
DB 2175 TMGDRIYVEIQLS---SIT-GSSSPACGASATICORANDOHFSRKVGTINQRIYVQDD 2230
QY 546 ---VFT-----VTPDTRSKVYLK-TPNMDRGLPSLTSSVMNISVPRDQVACLTFPE 593
DB 2231 LNVFTSSSCGCKDKTKSVSSTIFPHCDPLVKGDIPEFS-----HETADCCQLFW 2281
QY 594 RGVVCCQTRAMIIIOEQTRABEISLEEDVLPKSFHHHGFVNINISCSPTSGQLDL 653
DB 2282 HTSACVPLAGF-----DEETAGDDAOEHK-----GLSSRSQAVGAVLSL 2321
QY 654 LPSVTLPRTVLTVILIAVGGVLLSLALGHIIC-----YKKKKKTKNGPANGIYN 708
DB 2322 L-----LVALTACLITLLTKKRRRMMVSR--LTNCRSANSVYIKSVKKEBEADENE 2375
QY 709 XNINTE-MPROPKFKOKRKUNDSHVAV---IEDTM--VYGHLIQDSGSLQ-PEVD 760
DB 2376 TWMLEBIEIPAPRPEKEQE--NGHVAAKSVRAADTLALHGD-EQDSEDEVLTLPYVK 2432
QY 761 TTRPPO-GTMGVCPSPPTICSRAPTAK-----LATEPPRSPSESEPTTSHN 812
DB 2433 VRPFGAPGAEQ-GPELRFLPRKAPPLRADRVGLVRGEPARGRPRAAATPTSTFHD 2491
QY 813 NGD 815
DB 2492 SDE 2494

RESULT 6

T09456
Intrinsic factor-B12 receptor Cubilin precursor - human
C:Species: Homo sapiens (man)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #ext_change 09-Jul-2004
C:Accession: T09456
R:Kozyraki, R.; Kristiansen, M.; Silahatoglu, A.; Hansen, C.; Jacobsen, C.; Tommerup, N
Blood 91, 3593-3600, 1998
A>Title: The human intrinsic factor-vitamin B12 receptor, cubilin: Molecular characteriz
ion.
A:Reference number: Z16677; MUID:98241400; PMID:9572993
A:Accession: T09456
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3623 <K0Z>
A:Cross-references: UNIPROT:O60494; UNIPARC:UP10000070C4D; EMBL:AF034611; NID:93929528;
C:Genetics:
A:Map position: 10p12
C:Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology
C:Keywords: receptor; vitamin B12 uptake
F:1-24/Domain: signal sequence #status predicted <Sig>
F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>
F:436-467/Domain: EGF homology <EGF>

Query Match 2.7%; Score 117; DB 2; Length 3623;

Best Local Similarity 18.2%; Pred. No. 20;
Matches 110; Conservative 77; Mismatches 229; Indels 190; Gaps 26;

QY 52 LAKPCYIVSKHITMISIKGERI--VFTSCQSPENHVFIEIQNI-DCMSGCPFG- 107
DB 400 LSHPCD-----NGCIDTVGIGFYCKDSGWTGVNCTENINECLSNPCLING 445
QY 108 -----EVOLOPSTSLPTNRTF-----IWDVK 130
DB 446 TCVDDGVSFSCETRLMTGALCQVPOQVCGESISGINSFYSRSPDVGVHVNCEFWIK 505
QY 131 AHKSIGLELOFSIPLRLOIGPESCPDGYTHSISGRIDATVVRIGFCSNGTVSRIKMOE 190
DB 506 TEMGKVLRTITTFPRLESM---DNCHEFLQYVDDG-SSAPOLGFCSSLPHELSSD 561
QY 191 GYKMAHLHLPFHRNVSGFSIANRSSIKRLC--IISVPEGGSATLMSANYPEGPEDE 248
DB 562 NALVFHLYSEHLRNRGRTVWETO-QPECGILTGPY-----GSIKSPGYGNYPGR 614

QY 249 LMTQFVVPALRLASVSLFNFLSNCKERKEVEYIIPGSTNPEVKLED-----KOP-- 302
DB 615 DCWALVTVSPDLLVFTFTGTLSEHND-----DCKKRYLEIRDSPLQDPL 661
QY 303 GNMAGNFSLQCGDQDQSPG-IILQFOVLVQHONESNKIYVVDLSNERAMSLTIEP 361
DB 662 GKFTTFSV-----PILQTPPARIFHSDQISDQGFHTITLSPSDLRCGGNYPDP 715
QY 362 RPKVSKRFVP---GCFPLCEBRTSSNULTLTSCKKHSPLCDLTLMMNVEKTISCT 418
DB 716 ---GGEFLPELSPFT--HTRQCVVMKQDQGOIINF-----THVE----- 754
QY 419 DHRVQORSYSLOVPSDILHLEPELHDFMSKLLVPEKRLSLVLPDQKLQOHTHEKPCNT 478
DB 755 -----LQCCSDSSQNTIEVRD----- 770
QY 479 SFSYLVASAIPSQDLFSGFCPGSIIKOIOVKONISVTLRTFAPSPKOEAS-ROGLTVSF 537
DB 771 -----GETLLGVCGNGTISHIKITN--SWIRFKIDASVEKASFRAYQVAC 817
QY 538 IIFYKEGVFTYTPDTRSKVYIKRTNMDRGLPSLTSSVMNISVPRDQVACLTF--FKES 595
DB 818 GDELTEGEV-----IRSPFPNVYPGERTCMTIHQPOSQVLLNFTVFEIGS 865
QY 596 GVVCOT 601
DB 866 SAHCET 871

RESULT 7

AH2011
heterocyst glycolipid synthase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #ext_change 09-Jul-2004
C:Accession: AH2011
R:Kazaki, T.; Nakamura, Y.; Wolke, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2011
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1263 <KUR>
A:Cross-references: UNIPROT:Q8W67; UNIPARC:UP100000CE16B; GB:BA000019; PIDN:BA078012.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: hglE

Query Match 2.6%; Score 113.5; DB 2; Length 1263;

Best Local Similarity 19.3%; Pred. No. 74;
Matches 170; Conservative 115; Mismatches 350; Indels 247; Gaps 42;

QY 79 TSCQSPENHVFIE-----IQK-NIDMSGCPFG-----EVOLOPSTSL 118
DB 417 TINVTQPNPKFIESPPYINTEPWIQNGDIPRAAGVSSFGCTGYHVLBEYPTL 476
QY 119 P-----TLNRTIMDKAKHSIGLELOFSIPLRLOIGPESCPDGYTHSISGRIDAT 170
DB 477 PQGGRHSTAGSILLM-ADTPQQLSQCKQATLBQLQSDNGEGRYNOQJNHSKTLIPET 535
QY 171 VRIIGTFCNGTVSIIKQEGYKALHLP---WHPRNV---SGFSIANRSSIKRLCI 222
DB 536 SARVGFVATISLEBAQKLKALIKOLOIOPQAPAMHPOGIYVRQGTLSIKG----- 587
QY 223 IESVPEGGSATL-MSANYPEGPE-----DELMTQFVVPALRLASVSLFNFLSN 273
DB 588 VVALFPQGSQYLLNNGRELSLNFPEIQAYQALDQMA-QDNLPA-----LSD 634
QY 274 CERKEVEYIIPGSTNPEVKLEDKOPGNMAGNFNLSQCGDQDQSPGILRLQFOVL 333

```

Db      635 -----VFPIIP--AFSPETIKQSQQ-----LQRTENAPAIAGLSGLYKI 674
Qy      334 VQ-----HQNESNKIYVVDLSNERAMSLITERPKYQSKKFPGCGVCLSEST 382
Db      675 LQKVGFKPDPVAGHSGFGLTALMAAGVFSDEDEYCYLIKAR--GQAMTIPOG-----SHD 726
Qy      363 CSSNLTLL--TSGSKHKHSIFLCDDLTRLM-----NVEKTIISCTDHRX 422
Db      727 CDRCMTLAVSGDVAKIKOLIAGMTKQVQIANYNPAPEOVLSGSKPELALKEVLS-----780
Qy      423 CQRKSYSLQ--VPSDILHLPELHDPFSWMLVLPKRLSLVLVP--AQKLOQHTHEKPCNTS 479
Db      781 --KQGYTATPLTVSAFAHTPFVRH-----ASQPFMALRLVTFNTPQIVP 823
Qy      480 FSYLVASAIPSODLIFGSGFCPGSGIKQIVKQNIISVTLTAPSPFXQEASRQGL-----533
Db      824 YANMTGNAYTEATATIRQLLEAHLNLAVOFAQEIENLVAQGYCFVEIGPROILITNLVKQ 883
Qy      534 TVSFIPYFKEEGVFVTVPDTK--SKVYLR-----TPNMDRGLPSLTS 573
Db      884 TLGDPRHT---AIALNPSRQKSDVQLKQGVYQLRWGLSLSEDDASP--RKLPAQKS 936
Qy      574 VSMNISVPRDQVACLTFFKERSGVWCOTGRAF-----MIOEQRTRAEELFSLDEVL 626
Db      937 KSKGSLI---KLMTATNYISDK-----TKAFAEFAPKEPVPVQLASTPLDQVADAAATV- 986
Qy      627 PKPSTHHHSFWNINISNCSPTSGKOLDLFSVTLTPRTV-DLVYLIALAVGGVLLLSALG 685
Db      987 -----VSDSPAQPLVASNSND--EIAITPVSAVESVSL-----GEIYHQSDYQ 1027
Qy      686 LIICVKKKKKKTKNGKPAVGIYXNINTEMPPRQPKFOKGRD---NDSHYAVLIEDTM 741
Db      1028 LLI-----SKDET-----TMNSTHNTVAVLFOQFYHQKEMQLQHEQYVKTQSQSQ 1076
Qy      742 VYGHLLQDSSGSLQPEVDVYRPFQGTMCV-----CPSPPTICSRAPTAFLATEEPP 795
Db      1077 AFLQLLQOQ--EVLIPQGSTSAIEQTVVPAVVELPQTPVAVSVELPKTVAVPVESE 1134
Qy      796 RSPSESESEPTFFSHPNN--GVVSSKDTDIPILXTOEPNEPA 835
Db      1135 KVTAPVAPVAPHAISHANLEBRYSAPFPQPLVVTETASPA 1176

RESULT 8
Species: Mus musculus (house mouse)
C.Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C.Accession: S50065
R.Crocker, P.R.; Mucklow, S.; Bouckson, V.; McWilliam, A.; Willis, A.C.; Gordon, S.; Mill
EMBO J. 13, 4490-4503, 1994
A.Title: Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells
A.Reference number: S50065; MUID:95009950; PMID:7925291
A.Accession: S50065
A.Molecule type: mRNA
A.Residues: 1-1694 <CRO>
A.Cross-references: UNIPROT:O62230; UNIPARC:UP10000028B74; EMBL:Z36293; NID:G557253; PIR
Query Match      2.6% Score 113; DB 2; Length 1694;
Best Local Similarity 20.6%; Pred. No. 13;
Matches 109; Conservative 64; Mismatches 191; Indels 166; Gaps 25;

230 EGSATLMSANYPEGPEDELMTW-----QFVPAHLR--ASVSFLNFLNSCKERKERV 281
Db      903 EGOAVTLSCOVPTGVSEGTYSWYQDGRPLQESTSTLITAIASLRQAGAYHCQAQAPDT 962
Qy      282 EX-----YIPGSTTNPVFKLEDKPGMMAGNPN-----LS 312
Db      963 AIASLAAPVSLHVSSTPRHTVLSALLSTDPERLGHVCSVQSDPPAQQLFHRNRLVAST 1022
Qy      313 LOGCDQDQAS-----PGILRLQFVLVQHPQNESNKIYVVDLSN-----351

```

```

Db      1023 LQGADELAGSNRLHVTVLPNELRLQ-----IHPPELEDGCTVTCESASNTLQGAASAADPD 1078
Qy      332 ERAMSLTTEPRPVKSKRPVPGCFVCLASRTCSSNLTLTSGSKHSIFLCDDLTLM---408
Db      1079 AQAVRVTVPNATVQEGQV-----NLTLCLVWSTHQ-----SLSYTWYWG 1119
Qy      409 -----MNVEKTIISCTDHRXC-----QKRSYSIQVPSDILHLPELHDPFSWML 451
Db      1120 GQQLLGARSITLPSVKVDANISYR--CGVGLRGHAPHLSPVTLVDVLAHPRNLR-LTYLLE 1177
Qy      452 VPKRLSLVLV-----PAOKLOQHTHEKPCNTSFSYLVA---AIP-----489
Db      1178 TQGRQLALVLCTVDSRPPAQQLTSLHGDQ-----LVASTEASVPTNLALRELQDPRP 1228
Qy      490 -SQDLY-FGSGCPGGS-----IKQIQVKONISVTLRTAPSPFXQEASRQGLTVSFTP 539
Db      1229 SNEGLYSCSASHPICKANTSLLELLEGVRVKNPFGSVDEGEPVTVTCEDPAALSSALYA 1288
Qy      540 YFK-----EEG-----VTVTPDRTKSKYLRTPNMDRGLPSLTSVSMNI-SVPRQVACL 588
Db      1289 WFHNGHMLQEGPASSLQPLVTTRAHAGAYFCQVHDQTGRSSRPASLOILVAPRD--AVL 1346
Qy      589 TFFKERSGVWCOTGRAFMIQ---EQRTRAEELFSLDEVLPKPSEFHHHS 635
Db      1347 SSFRD-----SRTLMVVIQCTVDSERPAPAEVVLSHNKKVLAASHERHSS 1390

RESULT 9
Species: Homo sapiens (man)
C.Date: 13-Sep-1991 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C.Accession: A36688; A34210; A3613; A30187
R.Mikol, D.D.; Alexakos, M.J.; Bayley, C.A.; Lemons, R.S.; Le Beau, M.M.; Stefansson, K
J. Cell Biol. 111, 2673-2679, 1990
A.Title: Structure and chromosomal localization of the gene for the oligodendrocyte-my
A.Reference number: A36688; MUID:91115558; PMID:2277079
A.Accession: A36688
A.Molecule type: DNA
A.Residues: 1-440 <MI>
A.Cross-references: UNIPROT:P23515; UNIPARC:UP10000130CE1; GB:X57436; NID:G23105; PIDN
R.Mikol, D.D.; Gulcher, J.R.; Stefansson, K.
J. Cell Biol. 110, 471-479, 1990
A.Title: The oligodendrocyte-myelin glycoprotein belongs to a distinct family of protei
A.Reference number: A34210; MUID:90130636; PMID:1688857
A.Accession: A34210
A.Molecule type: DNA
A.Residues: 8-20, 'X', 22-440 <MI2>
A.Cross-references: UNIPARC:UP1000016AE0D; GB:X51694; NID:G35144; PIDN:CAA35991.1; PID
R.Vibkochil, D.; Cawthon, R.; O'Connell, P.; Xu, G.; Stevens, J.; Culver, M.; Carey, J
Mol. Cell. Biol. 11, 906-912, 1991
A.Title: The gene encoding the oligodendrocyte-myelin glycoprotein is embedded within t
A.Reference number: A39613; MUID:91117257; PMID:1899288
A.Accession: A39613
A.Molecule type: mRNA
A.Residues: 1-20, 'X', 22-440 <VIS>
A.Cross-references: UNIPARC:UP10000163BC8; GB:M63623; NID:G189385; PIDN:AAA59970.1; PIR
R.Mikol, D.D.; Stefansson, K.
J. Cell Biol. 106, 1273-1279, 1988
A.Title: A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in centra
A.Reference number: A30187; MUID:88198371; PMID:3283151
A.Accession: A30187
A.Molecule type: Protein
A.Residues: 25-29, 'H', 31, 'H', 33-44, 'D', 46-50, 'P', 52 <MI3>
A.Cross-references: UNIPARC:UP1000017C2D8
C.Genetics:
A.Gene: GDB:OMG; OMGP
A.Cross-references: GDB:127563; OMIM:164345
A.Map position: 17q11.2-17q11.2
A.Introns: #status absent
A.Note: this gene lies within an intron of GDB:NFI on the opposite strand
C.Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linka
P.1-24/Domain: signal sequence #status predicted <SIG>

```

F:56-78/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
F:79-99/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
F:100-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F:124-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
F:147-167/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
F:168-191/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
F:192-215/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
F:216-227/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>
F:426-440/Domain: carboxyl-terminal propeptide #status predicted end (Ctp)
F:425/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Aen) (in mature form)

Query Match 2.6%; Score 112.5; DB 2; Length 440;
Best Local Similarity 19.1%; Pred. No. 1.8;
Matches 79; Conservative 62; Mismatches 144; Indels 129; Gaps 17;

QY 256 VPAHLRASYSLNF-----NLSNCRKEERYE--YIPGSTTPEV-- 294
DB 49 LPSGLQENITHNLSTYHFTDLHNOLOYNLRTLLDISNNRLESLPAHLPRSLIMNSAAN 108
QY 295 --FKLEDKOPGNMAGNFNLSLQCDOD-----AOSPGLRLFOVLVQHPONES 341
DB 109 NNKIKLDR--SDPAVWMLKYLIDVSGKMLEKVVLIKNTLRSLLEVNLSSNKLMTYPTNMP 166
QY 342 NKIYVVDLSNERAMSLTIEPRPYKGRKEVPGCFVCLSERTCSNLTLTSGSKHKISFLC 401
DB 167 SKLHIYDLNN--SLT-----QLPGLTL-----NLT----- 191
QY 402 DDLTLRMMNVEKTICTDHRRCORRYSIQVPSDILHLVELHDFSEKLLVPRKRLSLVL 461
DB 192 -NLTLHLNHNKFTFLIPDSFDQ--LFOLO-----BITLYNNWS--CDHKONTYLL 239
QY 462 VPAOKLQGHTEKPCNTSFYSYLA--SAIPSDLYFSGFCGSGIKOIQVONISVTLR-- 518
DB 240 KMMETKAVIGIPCGTQJSSLKEHNMVTPPGFTSSLTPVSGMQVVDVINSLSVTQPK 299
QY 519 -----TPASFXQASRQGLTVSPFY-----FKEE 544
DB 300 VTKIPQVREKTEFTFGATLSKDTFTSTDKAFVPEDEDTSTETINSHAAATLTILHLD 359
QY 545 GVFET--VTPDTRSKKYTLATPMMDRLPSLTSVSNMISIPROVACLTFFKERS 595
DB 360 GNVNTLSLSTKSSPTPMTLSTTSGMN-----NFSEMQOSTTLNLMREBT 407

RESULT 10

VCDAR

env polypeptide - avian reticuloendotheliosis virus

N:Alternate names: coat polypeptide

N:Contains: coat protein gp22; coat protein gp73

C:Species: avian reticuloendotheliosis virus

C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004

C:Accession: A03999

R:Wilhelmsen, K.C.; Eggleston, K.; Temin, H.M.

J. Virol. 52, 172-182, 1984

A:Title: Nucleotide acid sequences of the oncogene v-rel in reticuloendotheliosis virus

A:Reference number: A93003; MUID:85009850; PMID:6090694

A:Accession: A03999

A:Molecule type: DNA

A:Residues: 1-562 <MIL>

A:Cross-references: UNIPROT:P03399; UNIPARC:UP10000129FE4

A:Experimental source: strain A

A>Note: strain A is a helper virus of the strain T

C:Comment: Enzymatic cleavages of env polypeptide may yield mature proteins including C

C:Genetics:

A:Gene: env

C:Superfamily: type C retrovirus env polypeptide

C:Keywords: coat protein; polypeptide

Query Match 2.6%; Score 112; DB 1; Length 582;
Best Local Similarity 19.1%; Pred. No. 3;
Matches 89; Conservative 59; Mismatches 148; Indels 170; Gaps 19;

QY 314 QGCDODAGSGILRLQFOVL--VQHPONESNKIYVDL-----SNERAMSLTIEPRPVK 365

DB 221 RGVDDPQTSIDLKTHQVLNATNQLENMCLMCTLTGSPQSPRRMAMSLSM----- 275
QY 366 QSRKRVPCGFCVLESRTSSNLTLTSGSKHKIS---FLC---DDLRLMMNVEKTICTD 419
DB 276 -----IAVLASLSGATHRVNRCOLLCREADNRTGIPGVGHFTNCTS 317
QY 420 HRYCQRKYSLOVPSD-----ILHLPEVLDPSKMLVPRKRLSLVLPQAOLQGHTEK 474
DB 318 IQESLTRRYIEILDVYLHRVYLCVEQHAAT--ALPNKKTGL----- 359
QY 475 PCNTSFYSYLAIPSDLYFSGFCGSGIKOIQVONISVTLRTFAPSFXQASRQGLT 534
DB 360 -C-----ILASIVDWM-----SIRPG-----EPIPL-----PSIETARHARA 393
QY 535 VSPIFYEEBEGVFTY----- 556
DB 394 VQFIFLVLGLGITGATLGGTGLGVSVHTYHKLNSQLIEDVQALSGTINDLDQDLSLAE 453
QY 557 VYLRTPMNDRLPSLTSVSNMISVPRDQACLTFFKERSGVVCCGGRAFMIIQERTAE 616
DB 454 VVLQN--RRGLDLTLAQGGICLALQKCC--FYANKSGIV-----RD 492
QY 617 EIFSLEDEVLK--KPSFHHHSFWVNISNCSPTSGKOLDLFSVTLTPRTVDLTLIAAVG 675
DB 493 KIRKIQEDLLAKRKALYNPLMNGLNGFLPYLPSLGLF----- 532
QY 676 GGVLLSLALGLIICCVKKKKKTKNGPRAVGYNNKINTEMROPK 721
DB 533 -GLILFLTLTG--PCIRKTLTRIHDKIQSKNRPISPAVATPMR 574

RESULT 11

hypothetical protein YHR155w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004

C:Accession: S46754

R:Macri, C.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of S. cerevisiae cosmid 9666.

A:Reference number: S46752

A:Accession: S46754

A:Molecule type: DNA

A:Residues: 1-1228 <MAC>

A:Cross-references: UNIPROT:P38851; UNIPARC:UP1000013B2E7; EMBL:U10397; NID:G500647; PID

A:Gene: MIPS:YHR155w

A:Cross-references: SGD:S0001198

A:Map position: 8R

C:Superfamily: yeast SIR3 protein; pleckstrin repeat homology

Query Match 2.6%; Score 112; DB 2; Length 1228;
Best Local Similarity 18.8%; Pred. No. 9.2;
Matches 99; Conservative 84; Mismatches 197; Indels 148; Gaps 21;

QY 179 SNGYTSRKIMQBGVMAHLPMFHRANVSGSIANKRSIKRLCTIES-----VFEEGAT 234
DB 301 SNGNNSSEKGLSG-----MLYKKTIVGDPKRVVWRWCFLQNNVFGVFSLSPEKT 352
QY 235 LMSANYPGCFPEDELMTQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTPEV 294
DB 353 YVEETDKG-----ILMTY-----EYLP----- 371
QY 295 FKLEDKOPGNMAGNFNLSLQCDODAGSGILRLQFOVLVQHPONESNKIYVDLSNERA 354
DB 372 -----KEBRNFCFKRLIQNPCKTEBENTYI-----DILLQAESIDELKSWINTLSHR 421
QY 355 MSLT-----EPRPVKQSKFPVCGVCLSERTCSNLTLTSGSKHKISFLDDELRLMMN 410
DB 422 IALSTKEENDRYOLARRKIEPQFEFASSTSTDKLTLSFSSKTLT-LVEELKNYMS 480
QY 411 VEKTICTDHRVQCRKYSYLOVPSDILHLPEVLDPSKMLVPRKRLSLVLPQAOLQGHTEK 454

Db 481 EDDIISIIDNKAYHARVSTPIATQTLHAL-----FSTFLSNTNPPCATQANTGTAMW 536
Qy 455 DRLSLVLPVPAOKLQOHTHEKPCNTSFYLVASAISS-----ODLYGFSFCPGSGSIKQ 506
Db 537 NDLSYLVNPLKSSVH---KPAIVNSSRFSVSDYYPYSLKVDIDQIFRSIFSVNDF 593
Qy 507 IOYKONISVTLR---TFAPSPYQF-ASROGLTVSIFPFKEGVTVP-----DYKSV 557
Db 594 LQVPEKL-VLLRYSSVWCPCNNKQKFSMAFVTLNHYLVNIGFSYLRRIDLDDISIE 652
Qy 558 YLRTP-----NMDRGLPSLTYSWNISV-----PRDVA 586
Db 653 YDKSKXHYSSRMHLNMQG-----DGLRFMSVFTTRRAVASKLOFLIENKAMHLPKGEKE 708
Qy 587 CLTFPKERSGVVQCOTGRAFMIIQEQRTABEIFSLDEVDLPKPSFHH 634
Db 709 VLEIFQELDEETENEK---IHKDNLSESEH-YSKDYDYLKSTYDHH 752

RESULT 12

J01674

protein kinase TMK1 (BC 2.7.1.-), receptor type precursor - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004

C/Accession: J01674

R/Chang, C.; Schaller, G.E.; Patterson, S.E.; Kwok, S.F.; Meyerowitz, E.M.; Bleecker, A.

Plant Cell 4, 1263-1271, 1992

A/Title: The TMK1 gene from Arabidopsis codes for a protein with structural and biochemical

A/Reference number: J01674; MUID:93076110; PMID:1332795

A/Accession: J01674

A/Molecule type: DNA

A/Residues: 1-942 <CH>

A/Cross-references: UNIPROT:P43398; UNIPARC:UPI000048404; GB:I00670; NID:G166887; PIDN:

C/Superfamily: Receptor-like protein kinase; leucine-rich alpha-2-glycoprotein repeat hc

C/Keywords: ATP; autophosphorylation; glycoprotein; phosphotransferase; receptor; serine

F/1-22/Domain: signal sequence #status predicted <SIG>

F/23-942/Product: protein kinase TMK1 #status predicted <MAT>

F/65-88/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F/89-111/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F/112-135/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F/136-160/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F/161-186/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F/187-209/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F/210-232/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F/233-255/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F/256-279/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F/280-299/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F/300-323/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F/324-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F/363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F/387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F/411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F/440-503/Domain: transmembrane #status predicted <TM>

F/566-872/Domain: protein kinase homology <KIN>

F/599-602/Region: protein kinase ATP-binding motif

F/86-99/158,161,171,353,533,587/Binding site: carbohydrate (Asn) (covalent) #status pred

F/616,634,717,719/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 2.5%; Score 111.5; DB 1; Length 942;

Best Local Similarity 19.4%; Pred. No. 6.7; Indels 233; Gaps 29;

Matches 128; Conservative 82; Mismatches 216; Indels 233; Gaps 29;

Qy 117 LPLPLNRTFWDKAHNSIGLEQFSIPRLRQIGSGSCPDGVTSHI-SGRIDATVVRIG 175

Db 16 LLLSLTKR---DSQGLSAMLSTLKSLSNPSPSFGMSDDPCKMTHIVCTGRKVTRIQIG 72

Qy 176 TFCNGYVSRKIMQEGVMAHLPMFHRPNVSG--FSJANSSIKRLCLTISVBEG---- 229

Db 73 HSGIQTLSPLRLNLSLELERLEQW--NNISGVPVSLSGLASIQVLMLSNNNPDSPSD 129

Qy 230 --EGSATLSANYPEGFPEDELMTQFVVPALRLASVSLFNFNLSNCRKEKREVEYIIPG 287

Db 130 VFOGLTSLQSV-----IDNPFKSMW--IPESLNAASLQNFSAASA-----NVSGSLPG 178
Qy 288 STNPEVFKLEDKQPG-----NMAGNENISLQCCDODAQSPGILRLQFOVLVQHPO 338
Db 179 -----FLGDEFFPGSLILHAFNNLEGEFLMWSLAGS-----QVOSLMLNGQ 219
Qy 339 NESKIIYVVDLSNERRASLTIEPRPVKOSRKFVGCFCPLCESRCCSSULTLTSQKHS 398
Db 220 KLTDITV--LQNTGTLK-----EYWLHNSNKRSGPLPDFSGIKEL-- 257
Qy 399 FLCDLTRLMMNVKETSCTDHRVCQKRSYSLQVPSDILH-----PYEL 443
Db 258 -----ESLSLRN-----SFTGPPASLISLSLKVNVTNNHLAGPVPV 297
Qy 444 HDFSMLKLVPRK-----LSLVLV-----PAOKLQOHTHEKPCNTSPSY 482
Db 298 FKSVSVDLDKDSNFCILSPGECDPVKSLILLASSPDYPRLAESWKGNDPCTN----- 353
Qy 483 LVASAISSQDIYPSFCPGSGSIKQIOYKONISVTLRTPAPSPXOBASRQGLTV-----S 536
Db 354 -----WIGIACSNGNITVLSL-EKKMELT-GTISPEFGAIKSLQRIILGINNLTG 400
Qy 537 FIPYFKEGVFTVPDRK-----SKVYLRTPNMDRGLPSLTYSWNISVPRDVAULTF 590
Db 401 MIPQ-----ELTTPNLKTLDVSSNKLPGKVPGF----- 429
Qy 591 FKERSGVVQCOTGRAFMIIQEQRTABEIFSLDEVDLPKPSFHHHSFWNINSCSPTSG-- 648
Db 430 ---RSNVVV-----NTNGNDPIDGDKSSLSF-----GSSPSGSG 463
Qy 649 -----KQDLDFSVTLTPRVVDLTVLLIAVGGVLLSMAGLITCVKKKKKTKMG 701
Db 464 SGINGDGRGKMSSTPIG-----IIVGSLVGLSLIFLIGLVPCWYKROKRFSG 515

RESULT 13

hypothetical protein T01G9.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T24315

R/Lienhard, N.

submitted to the EMBL Data Library, July 1996

A/Reference number: Z19874

A/Accession: T24315

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-603 <MTL>

A/Cross-references: UNIPROT:Q22075; UNIPARC:UPI00000762F8; EMBL:Z75713; PIDN:CA800050.1

A/Experimental source: clone T01G9

C/Genetics:

A/Gene: CESP:T01G9.3

A/Map position: 1

A/Introns: 68/3; 346/1; 407/3; 450/3; 486/3

Query Match 2.5%; Score 111; DB 2; Length 603;

Best Local Similarity 19.7%; Pred. No. 3.8;

Matches 132; Conservative 79; Mismatches 217; Indels 242; Gaps 32;

Qy 266 FLNF-----NLNCRKEKREVEYIPGSTNPEV-----KLEDK 300

Db 9 FLSLFLVLCIALPSSCNLCEDONDSSWSVCKAIINTTIAEILINQPLTLRSLHIQ 68

Qy 301 QPGNMAG-----NFN-----LSLQCCDDADQSPGILRLQFOVLVQHPO-----E 340

Db 69 PPSRIGSNKLNKNNNDINRRAQLRVLRILNCOIPAMRSRIRLPSLEVLDSHNNIEHATM 128

Qy 341 SN-----KIYVVDLSNERRASLTIEPRPVKOSRKFVGCFCPLCESRCCSSULTLTSQKHS 393

Db 129 SNFGMKLKVLDLSSNH--LNILP-----TGVTYLRALRSLSLSNNTTIDLS 175

Qy 394 KHKISFL-----CODLTRLMMNVKETSCTDHRV---CQKRS-YSLQVPSD 435

Db 176 TNLRLGLNSLRVLRDRPIPIEHINELFTDY-----SQDLDELXLANHCLNLSISYSLADR- 230
Qy 436 ILHLPEVLHDSKWLKLVKRDLE-----SLVLYVPAQKLOQHTHEKPCNTSPSYLVASALP 489
Db 231 -----IPQLRQLGIGNNLKAVPTKELR-----LPQSLVDLDSHNS 267
Qy 490 SODLYFGSPCGSJKQIQVKONI-----SVTLRTFAFSF----- 524
Db 268 IQEIRACAF-C-NTNISKDLSHNLGISKDSPFNEDAPRTMPLRLDLSPNMDPDSKW 326
Qy 525 ---XQASRQGLTVSFIPYKEEGFTVTPDTKSKVYLRT-PNDRGLP----- 569
Db 327 LGMAQEEELTSLALSGNFKNFEESEWYTL-----KSLIHLELAYNHKIFPVQLPARYHL 382
Qy 570 -----SLTSVSNISVPRDQVACLTFFKERSGVVCTGAPMIIQEQ----- 611
Db 383 ISLINSIGNELTYLPNINITLLPNVKTFTDITANRFTHTSHTDLAFANNEQVYVDGNPMD 442
Qy 612 -----RTRAEIFSLDEDLVLPKPSF-----HHHSFWVNISNCSPTSGKOLDLLE 655
Db 443 SCATGLOVHMRDRVAMRHILNYDVNRCATPSLVESHSLAITDVNDCAVLFGARYGL-- 500
Qy 656 SVTLPTPTVDLTVILIAAVGGVLLLSALGLIIC-CVKKKKKTKNKGPVGIYNNXNITE 714
Db 501 ---TQISEMILIL-----AGVLLFALLMLITACIYFLRROYKG-----SYVTR 543
Qy 715 MPQPKKKQKGRKNDSHVYAVIEDTMYGHLLODSSGSF---LQPEVDTY---RPRQG 767
Db 544 HSRITL-----TMANTHSCSSSTMDTHGRLPDPDPFLVSTETFXA 584
Qy 768 TMGVCPSPP 777
Db 585 TPRLTPPAP 594

RESULT 14

T19212

hypothetical protein C12C8.3 - Caenorhabditis elegans

C.Species: Caenorhabditis elegans

C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C.Accession: T19212

C.Barlow, K.

submitted to the EMBL Data Library, November 1996

A.Reference number: Z19091

A.Accession: T19212

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-1185 <WIL>

A.Cross-references: UNIPARC:UPI000017B772; EMBL:Z81467; PIDN:CA803872.1; GSPDB:GN00019;

A.Experimental source: clone C12C8

C.Genetics:

A.Gene: CESP:C12C8.3

A.Map position: 1

A.Introns: 46/1; 133/2; 150/1; 197/1; 287/3; 378/2; 419/3; 449/1; 553/3; 580/2; 704/2; 7

Query Match 2.5%; Score 111; DB 2; Length 1185;

Best Local Similarity 19.0%; Pred. No. 10; Mismatches 312; Indels 346; Gaps 48;

Matches 181; Conservative 114; Mismatches 312; Indels 346; Gaps 48;

Qy 17 LUGAARLPGRGAFAEIALPRESNITVLIKLGTPTL-LAKPCYIVISKRHITMLSIKSGER 75
Db 310 LMASPDVPR--HSATIAPRNS-----MCSIPRLQLATP-----MSGSGSQ- 348
Qy 76 IVFTSCSGSPENHFVIEIQKIDCMGCPPEBVOLOPST-----LLPTLNRTFIMDV 129
Db 349 ---TTPISBPLSGOPQOO-----PMGPIDQCGESKISFAYVQCCEALCICV 395
Qy 130 KAHKSIGLELPISIRLRQ-----IGPESCDGVTHTISGRIDATVVRIGFCGN 180
Db 396 QHQRVATKQHAFFELQOLMTLMSRAVQPOOA---QYTQVNGSVROALSGVGFESG 453
Qy 181 GTVSRIKMGEGVMALHLPLWFPFRNVSGFSIANRSSIKR---LC-----IISVEGEGS 232

Db 454 -----HVGENDSISGESSSPRSSVCGTHDSVIIGICENCPH 492
Qy 223 ATMS-----ANYP-----EGFPDELMTQFVVPAHLRASVFLNPLLS 272
Db 493 SVLLCAICVAHQPGKRRVQPLGDIRAVAGVEVNSQLQWO----- 533
Qy 273 NCEKREVEEYIIGSTTN-----PEVFKLEDRQPC-----NNAGFNLT--- 311
Db 534 -CEKTGDITIKIIGIYVNTAANEIRAAFDTHVNLAEERKELKLVETVKKLSVL 592
Qy 312 ---SLQGCDDAQ-----SPGLRLFOVLYVQ-----HPONESNK 343
Db 593 ISQAESLQSKOIIDQALQIATKLMSSDCDEMYLRQVFXFKIACQMGNEGTEPNN--NI 650
Qy 344 IYVVDLS---NEBMSLITEPRP---VKQSKRFVPGCVCLSERTCSNLLTSGSKHKI 397
Db 651 LNVMLAQVNEBDRLEKTAPODGLLNKAKQFN---IESGPCANSSIVGDSFKKI 705
Qy 398 SFLCDLTRLMMNVEKTICTDHRV-----CORKSYSLOVPSD 435
Db 706 -----INXKMKNS--RYSNHFLSQFELNFCIFQONQSEVCGKKNIAIREQT 752
Qy 436 ILHLPEVLHDSKWLKLE---VPKRLSLVLPQKLOQHTHEKPCNTSPSYLVASALPSQ 491
Db 753 VIY--VQLRDACGDLSSSIATOPTSQALLPHQ--EPHSH-----LEQAMPTS 797
Qy 492 DLYGSPFGGSIKQIQVKONI SVTLRTFAFSFQOASRQGLTVSFIPYKEEGFTVTP 551
Db 798 DVQAFVISPDSSTVEV-----TMTP-----RENGIVALSYPSIEGSYTLNT 839
Qy 552 DTKSKVYLRTPNMWRGLPSLTSGVSNISVPRDQVAC---LTFKERSG-----VVC 599
Db 840 LVKG-----TP--ISGCTTMDIRGRNY--DEIAKAPILTFGSESGDDELCPWIC 890
Qy 600 QTGAFMIIOBORTAEIIFSLDEDLVLPKPSFHHSFWVNISNCSPTSGK---QLDLFS 656
Db 891 VDQGRVIVADRSNNRVQIFDXDGNFISK-----FGTSGNPPQGFDRPAG 935
Qy 657 VTLPTPTVDLTVILIAVNG-----GVLILSALGLIICVKKKKKTKNKGPVGIY 707
Db 936 IT---TNSLNNIVADKXNHRVOVFDENGMFL-----KFGDRGRAVGYF 977
Qy 708 NXNINTEMPQPKKQKGRKNDSHVYAVIEDTMYGHLLODSSGSFLQ---PEVDTYR 763
Db 978 N-----YPMGVATN--SHNALVADTNRHNRVQITPQOQFVRKCGFDSAYPFK 1023
Qy 764 PPOGTWVC-PSPPTTCSRAPYAKLATBEPPRPSPESEBPTTFSHPNNGD 815
Db 1024 NLDSFRGLCYLPDGLLITDFNNHRLAV-----LSPNMSEMKVYGSBGDD 1070

RESULT 15

T42205

breast cancer susceptibility protein BRCA2 - mouse

C.Species: Mus musculus (house mouse)

C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Mar-2004

C.Accession: T42205

R.McCallister, K.A.; Haugen-Strano, A.; Hagevik, S.; Collins, N.K.; Brownlee, H.; Futreal

submitted to the EMBL Data Library, February 1997

A.Description: Characterization of the mouse and rat homologs of the BRCA2 breast cancer

A.Reference number: Z22073

A.Accession: T42205

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-3329 <MCA>

A.Cross-references: UNIPARC:UPI0000161CCC; EMBL:U89652; NID:G2443438; PID:G2443439; PIDN

A.Experimental source: strain CD1; 129Sv; ICR Swiss

C.Genetics:

A.Gene: BRCA2

C.Superfamily: DNA recombination repair protein, BRCA2 type

Query Match 2.5%; Score 111; DB 2; Length 3329;
Best Local Similarity 18.8%; Pred. No. 49;

QY	198	LPWFFRNVSGF-----SIANSSIXICLIIESVDEGSSATLMANYPGEDEBL	243
Db	1188	LPNINEMFEGGFCMALGKTLKSVSNALAKKAKMLPSDIENSEBPSAKVC--PROGSSS--	1242
QY	250	MTWQFVPAHLRAVSFLNFINLSNCER---KEEREVEYYIPGS--TT-----NPEVFK	296
Db	1243	-----AHHDSVASVAFKIKQNTKEKFEDEKSSKCVTLQNNIEMTTCIFVGRNPEKXI	1294
QY	297	LEDKOPGMAGNPNFLSLOGCDODDQSPGILRLQFVUWHQONENKTYV-----VDLSNE	352
Db	1295	KATKHEDSYTSSQRRNNLENSDQSSS-----TSGPAYIHKGSDL-----	1334
QY	353	RAMSLTIEPRPVKOSRKKEFPGCFVCLSERTCSSNLTLSGSHKIKSPICDDLTRL-VMNV	411
Db	1335	-----PADQSSK-----CPRECTQYARENQ-IKENIS-----DUTCEIENKA	1372
QY	412	EKITSCTDHRICQKRSYSLOVPSPDLHLPEVLHDFSM-----	448
Db	1373	EET-----CMKSSDKKOLPSPDKMEONIKEFNISFQTASGKNTRVASKESLNKSUNIFN	1424
QY	449	-----KLIVPKDRL-SLIVLPAPKLOOHT--HEKPCNTSFYLVASAIPODLYFCGFCPG	501
Db	1425	RETELTVIISLSINSKILHGINKDMKMTSCHKK-----AIIKIVFEBHF-PI	1471
QY	502	GSIKIOVAKONISVTLRTFPAPSFQOASRQGLTVSF-----IPYKX-EGVFTVTPD	552
Db	1472	VTVSOLPAQOCH-----PEYELESTKEPILLSFHTASGKVAKIMQESLDKYNLFPDE	1522
QY	553	TKSKVYLR-TEPNMDRGLSLTYSVSNISVPPDQVACLPFEKRSQVVCQTRAFMIIOEQ	611
Db	1523	TO--YVAKTISFSGSKPL-----KOSKELTLAYEK-----LEVT	1556
QY	612	RTABEET--ESLDEDLVLPKPSFHHHSFWANISNCSPTSGK-QLDLFSVTLTPRTVDLT	667
Db	1557	ASKCEEMQNFVSKETEMLPQONVHYMYROTENILKTSNGTSKVOEINENVEKNPR----	1611
QY	668	VILLAAGGVLLLSALGLIICVYKKKKKTKNKKGAVGIIYXNNINTENPRPKKFQKGRK	727
Db	1612	-----ICCCIQSSSYPVTEBALVYVEDSRKTCVRE-SLSISKRRK	1650
QY	728	-----DNDSHVAVIEDTMVY-GHLLODSSGSLQPEVDYVAPFGTGMVGCPS	775
Db	1651	WLREGQDKLGRNTIKIECVKHEHTEDEAFGNASYSERHSLVITINTEIDTNHVSNOVSTL-LS	1709
QY	776	PPTIC 780	
Db	1710	DPNVC 1714	

Search completed: May 4, 2006, 22:25:18
Job time : 50 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:12:25 ; Search time 119 Seconds

(without alignments)
4956.485 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377

Sequence: 1 MAGNCGVSIALGLVLLGA.....SSKTDIDPLKTFQEMEPAE 836

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80:*

1: uniprot_sprot:*
2: uniprot_trcprot:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4375	99.9	836	2	Q9H5V8_HUMAN
2	4374	99.9	836	2	Q96QU7_HUMAN
3	3629	82.9	697	2	Q6HT71_HUMAN
4	3603	82.3	691	2	Q6U9Y2_HUMAN
5	3574.5	81.7	833	2	Q5U462_MOUSE
6	3395	77.6	649	2	Q9H8C2_HUMAN
7	2041	46.6	392	2	Q9H676_HUMAN
8	1849.5	42.3	470	2	Q8B2D9_MOUSE
9	1779	40.6	343	2	Q8WU91_HUMAN
10	1579.5	36.1	384	2	Q921M9_MOUSE
11	828	18.9	867	2	Q4T223_TETNG
12	824	18.8	825	2	Q4T7U3_TETNG
13	630	14.4	160	2	Q810U0_MOUSE
14	610	13.9	793	2	Q58EM7_BRARE
15	141	3.2	746	2	Q01654_HALRO
16	140.5	3.2	888	2	Q5F10_ENTHI
17	140	3.2	887	2	Q6ONX4_CAEBR
18	139	3.2	746	2	Q81AD8_HALRO
19	137.5	3.1	719	2	Q9PVY2_TRISC
20	137.5	3.1	3623	2	Q70244_RAT
21	131.5	3.0	2796	1	CSMD3_MOUSE
22	127	2.9	595	1	SIGL1_HUMAN
23	126	2.9	832	2	Q91B4_MOUSE
24	125.5	2.9	508	2	Q18530_CAEBL
25	123.5	2.9	3620	2	Q9TUS3_CANPA
26	123.5	2.8	4060	2	Q6T3A5_MOUSE
27	122.5	2.8	421	2	Q7YFC4_CRIYP
28	121.5	2.8	2302	2	Q5F4U0_CAEBL
29	121	2.8	766	2	Q618F0_CAEBR
30	121	2.8	1242	2	Q4X2J9_PLACH
31	120.5	2.8	3570	2	Q70737_ANOGA

32	120	2.7	567	2	Q61DZ1_ANOGA	Q61d21 anopheles g
33	120	2.7	1516	2	Q504M1_XENTR	Q504m1 xenopus tro
34	119	2.7	597	1	SIGL1_PANTR	Q951h0 pan troglod
35	119	2.7	847	1	Q82J97_STRAW	Q82j97 streptomyc
36	119	2.7	941	2	Q4RY40_TETNG	Q4ry40 tetraodon n
37	119	2.7	2122	2	Q54R92_DICDI	Q54r92 dictyosteli
38	118.5	2.7	407	2	Q9D2J4_MOUSE	Q9d2j4 mus musculu
39	118.5	2.7	2476	1	ZAN_PIG	Q28983 mus ectofa
40	118.5	2.7	4059	2	Q8C1S7_MOUSE	Q8c1s7 mus musculu
41	118	2.7	944	2	Q6CSES_KLULA	Q6cses kluyveromyc
42	118	2.7	2520	2	Q6HFE0_BACHK	Q6hfe0 bacillus th
43	118	2.7	2766	1	PD2K3_RAT	Q9qzr8 rattus norv
44	117.5	2.7	819	2	Q747S7_GROSL	Q747s2 geobacter s
45	117.5	2.7	1424	2	Q91TRS_ARATH	Q91trs arabidopsis

ALIGNMENTS

RESULT 1	Q9H5V8_HUMAN	PREDIMINARY;	PRT;	836 AA.
ID	Q9H5V8_HUMAN			
AC	Q9H5V8;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-FEB-2005 (TREMBLrel. 29, Last annotation update)			
DE	Hypothetical protein FLJ22969 (NCSG135).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	NUCLEOTIDE SEQUENCE.			
RA	Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,			
RA	Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,			
RA	Nakamura Y., Isegaki T., Sugano S.,			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
RN	(2)			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=22547370; PubMed=12660814; DOI=10.1038/sj.onc.1206220;			
RA	Hooper J.D., Zijlstra A., Aimes R.T., Liang H., Claassen G.F.,			
RA	Tarin D., Teeta U.E., Quigley J.P.,			
RT	"Subtractive immunization using highly metastatic human tumor cells			
RT	identifies SIMA135/CDP1, a 135 kDa cell surface phosphorylated			
RL	glycoprotein antigen."			
DR	Oncogene 22:1783-1794(2003).			
DR	EMBL; AK026632; BAB15511.1; -; mRNA.			
DR	EMBL; AF468010; AAO33397.1; -; mRNA.			
SQ	SEQUENCE 836 AA; 92874 MW; 98980475C3E5C4C8 CRC64;			
Query Match	99.9%; Score 4375; DB 2; Length 836;			
Best Local Similarity	99.6%; Pred. No. 3.2e-314;			
Matches 833; Conservative	0; Mismatches 3; Indels 0; Gaps 0;			
Qy	1	MAGNCGVSIALGLVLLGAARLPARGAEFIALPRESNTIVLKLGTPTLLARCYIVI	60	
Db	1	MAGNCGVSIALGLVLLGAARLPARGAEFIALPRESNTIVLKLGTPTLLARCYIVI	60	
Qy	61	SKRHTIMLSKGRIVFTSCSPENHFVETQKNTDMSGPFPFGVQLQPSLSLPT	120	
Db	61	SKRHTIMLSKGRIVFTSCSPENHFVETQKNTDMSGPFPFGVQLQPSLSLPT	120	
Qy	121	LNRTFMDVKAKSIGELQFISRLRQIGGESCPGVTHSIGRIDATVVRIGTFCSN	180	
Db	121	LNRTFMDVKAKSIGELQFISRLRQIGGESCPGVTHSIGRIDATVVRIGTFCSN	180	
Qy	181	GTGVRITKQSGVKAHLPMFHPNVGFSIANSSTIKRLCTISVVEGEGSATLSANY	240	
Db	181	GTGVRITKQSGVKAHLPMFHPNVGFSIANSSTIKRLCTISVVEGEGSATLSANY	240	
Qy	241	PEGPEELMTQGVVVAHLRASVSFLNPNLSNCRKEERYEYVTPSTNPEVFXLEDK	300	
Db	241	PEGPEELMTQGVVVAHLRASVSFLNPNLSNCRKEERYEYVTPSTNPEVFXLEDK	300	

```

Db 241 PEGFEDELMTWQFVPAHLRASVSFLNFNLSNCRKERVEYY1PGSTTNPVEFKLBDK 300
Qy 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
Db 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 361 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVKPKRSLVLPVPAQKLOQTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVKPKRSLVLPVPAQKLOQTHEKPCNTSF 480
Qy 481 SYLVASAIPSODLYFGSCFCPGGSIKQIOVKONISVTLTLPAPSFQOASRQGLTVSFLPY 540
Db 481 SYLVASAIPSODLYFGSCFCPGGSIKQIOVKONISVTLTLPAPSFQOASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDTSKAVLYRTPNMDRGLPSLTYSVMNISVPRDQVACLTFKERSGVVQ 600
Db 541 FKEBGFVTPTDTSKAVLYRTPNMDRGLPSLTYSVMNISVPRDQVACLTFKERSGVVQ 600
Qy 601 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 601 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 661 PRVULVTYLLAAVGGVLLSALGLIICVKKKKKTKNKGPAVGIXNNINTEMPRPK 720
Db 661 PRVULVTYLLAAVGGVLLSALGLIICVKKKKKTKNKGPAVGIXNNINTEMPRPK 720
Qy 721 KFOGRKONDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPGTNGVCPSPPTIC 780
Db 721 KFOGRKONDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPGTNGVCPSPPTIC 780
Qy 781 SRAPPAKLATEBPPRSPPESESEBPTFSHPNNGDVSSKOTDIPLLXTOEPMEPAE 836
Db 781 SRAPPAKLATEBPPRSPPESESEBPTFSHPNNGDVSSKOTDIPLLXTOEPMEPAE 836

RESULT 2
Q96GUT_HUMAN PRELIMINARY; PRT; 836 AA.
ID Q96GUT_HUMAN PRELIMINARY; PRT; 836 AA.
AC Q96GUT;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CUB domain containing protein 1.
GN Name=CDCP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=2159860; PubMed=1146621; DOI=10.1038/sj.onc.1204566;
RA Scherl-Mostaguer M., Sommergruber W., Abseher R., Hauptmann R.,
RA Ambros P., Schwelger N., Sommergruber W., Abseher R., Hauptmann R.,
RT "Identification of a novel gene, CDCP1, overexpressed in human
RT colorectal cancer.";
RL Oncogene 20:4402-4408 (2001).
DR EMBL; AY026461; AAK02058.1; -; mRNA.
DR Ensembl; ENSG00000163814; Homo sapiens.
SQ SEQUENCE 836 AA, 92873 MW, PBA4D2DBBD35C519 CRC64;

```

```

Query Match 99.9%; Score 4374; DB 2; Length 836;
Best Local Similarity 99.6%; Pred. No. 3.7e-314;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 MAGNCGVSIALLGVLLGAARLPRAAEAFETALPRESNITVLIKGTTLAKPCYVI 60
Db 1 MAGNCGVSIALLGVLLGAARLPRAAEAFETALPRESNITVLIKGTTLAKPCYVI 60

```

```

Qy 61 SKRHITMLSIKGERIVTFSCQSPENHVLIEIQNIDCMGCPGFGVLOLPSTSLPT 120
Db 61 SKRHITMLSIKGERIVTFSCQSPENHVLIEIQNIDCMGCPGFGVLOLPSTSLPT 120
Qy 121 LNRFTIWDVKAHKSIGLELOPSIRLROIGGESCPODVYTHISGRIDATVYRIGTFCSN 180
Db 121 LNRFTIWDVKAHKSIGLELOPSIRLROIGGESCPODVYTHISGRIDATVYRIGTFCSN 180
Qy 181 GTVSRIKQOEVEKVALHLPWFHPRNVSGFSIANKSSIRGLCIISVFEGESATLMSANY 240
Db 181 GTVSRIKQOEVEKVALHLPWFHPRNVSGFSIANKSSIRGLCIISVFEGESATLMSANY 240
Qy 241 PEGFEDELMTWQFVPAHLRASVSFLNFNLSNCRKERVEYY1PGSTTNPVEFKLBDK 300
Db 241 PEGFEDELMTWQFVPAHLRASVSFLNFNLSNCRKERVEYY1PGSTTNPVEFKLBDK 300
Qy 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
Db 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 361 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVKPKRSLVLPVPAQKLOQTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVKPKRSLVLPVPAQKLOQTHEKPCNTSF 480
Qy 481 SYLVASAIPSODLYFGSCFCPGGSIKQIOVKONISVTLTLPAPSFQOASRQGLTVSFLPY 540
Db 481 SYLVASAIPSODLYFGSCFCPGGSIKQIOVKONISVTLTLPAPSFQOASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDTSKAVLYRTPNMDRGLPSLTYSVMNISVPRDQVACLTFKERSGVVQ 600
Db 541 FKEBGFVTPTDTSKAVLYRTPNMDRGLPSLTYSVMNISVPRDQVACLTFKERSGVVQ 600
Qy 601 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 601 PRPVAKSRKFPVGCFCVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 661 PRVULVTYLLAAVGGVLLSALGLIICVKKKKKTKNKGPAVGIXNNINTEMPRPK 720
Db 661 PRVULVTYLLAAVGGVLLSALGLIICVKKKKKTKNKGPAVGIXNNINTEMPRPK 720
Qy 721 KFOGRKONDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPGTNGVCPSPPTIC 780
Db 721 KFOGRKONDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPGTNGVCPSPPTIC 780
Qy 781 SRAPPAKLATEBPPRSPPESESEBPTFSHPNNGDVSSKOTDIPLLXTOEPMEPAE 836
Db 781 SRAPPAKLATEBPPRSPPESESEBPTFSHPNNGDVSSKOTDIPLLXTOEPMEPAE 836

RESULT 3
Q6NT71_HUMAN PRELIMINARY; PRT; 697 AA.
ID Q6NT71_HUMAN PRELIMINARY; PRT; 697 AA.
AC Q6NT71;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE CDCP1 protein (Fragment).
GN Name=CDCP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RP TISUP-Kidney.
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.U., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```

RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkin R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez R.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RG NIH MGC Project;
 RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC069254; AAH69254.1; -, mRNA.
 FT NON TER 697 697
 SQ SEQUENCE 697 AA; 77703 MW; 9180BB3FE0C3EB18 CRC64;

Query Match 82.9%; Score 3629; DB 2; Length 697;
 Best Local Similarity 99.7%; Pred. No. 3.4e-259;
 Matches 695; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGNCGSIALGLVLLGAARLPARGAEAFIALPRESNTIVLTKGPTLLAKCYVI 60
 DB 1 MAGNCGSIALGLVLLGAARLPARGAEAFIALPRESNTIVLTKGPTLLAKCYVI 60
 QY 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKIDMGSPCFEGVLOPSTSLPT 120
 DB 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKIDMGSPCFEGVLOPSTSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLELQSPRLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQSPRLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 QY 121 LNRFTIMDVKAHKSIGLELQSPRLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQSPRLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIRKQEGVKALHPWFHPRVNVSFSTIANSSIRKLCIISVFEGESATLMSANY 240
 DB 181 GTVSRIRKQEGVKALHPWFHPRVNVSFSTIANSSIRKLCIISVFEGESATLMSANY 240
 QY 241 PEGFPEDELMTQWVVPVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVFKLEDK 300
 DB 241 PEGFPEDELMTQWVVPVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVFKLEDK 300
 QY 301 QPGNAGFNLSLQGCDDQDQSPGILRLQFQVLVOHPONESNKIYVVDLSNERAMSLTIE 360
 DB 301 QPGNAGFNLSLQGCDDQDQSPGILRLQFQVLVOHPONESNKIYVVDLSNERAMSLTIE 360
 QY 361 PRPVQSRKFPVGCFCVLCIESRTCSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420
 DB 361 PRPVQSRKFPVGCFCVLCIESRTCSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420
 QY 421 RYCORKSYSLQVPSIILHLPVELHDFSWKLVPRKRLSLVLPACKOOHHEKCNMFSF 480
 DB 421 RYCORKSYSLQVPSIILHLPVELHDFSWKLVPRKRLSLVLPACKOOHHEKCNMFSF 480
 QY 481 SYLVASAIPODLVYSGFCPGSIRKQIQVKONISVTLRTFAPSFQOASROGLTYSFTPY 540
 DB 481 SYLVASAIPODLVYSGFCPGSIRKQIQVKONISVTLRTFAPSFQOASROGLTYSFTPY 540
 QY 541 FKESGVFTVPTKSKVYLRTPNMDRGLPSLTYSVSNISVPRDQVACLTFEKGSGVVCQ 600
 DB 541 FKESGVFTVPTKSKVYLRTPNMDRGLPSLTYSVSNISVPRDQVACLTFEKGSGVVCQ 600
 QY 601 TGRAFMIIQOORTAEIIFSLDEVDLPKSPFHHSFWNINSCSPSTSGKQDLFLFSVLT 660
 DB 601 TGRAFMIIQOORTAEIIFSLDEVDLPKSPFHHSFWNINSCSPSTSGKQDLFLFSVLT 660

DB 601 TGRAFMIIQOORTAEIIFSLDEVDLPKSPFHHSFWNINSCSPSTSGKQDLFLFSVLT 660
 QY 661 PRYDLYTILAAVGGGVLSSALGLTICCKKKKKK 697
 DB 661 PRYDLYTILAAVGGGVLSSALGLTICCKKKKKK 697

RESULT 4
 Q6U9Y2_HUMAN
 ID Q6U9Y2_HUMAN PRELIMINARY; PRT; 691 AA.
 AC Q6U9Y2;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Membrane glycoprotein gp140 (Fragment).
 GN Name=CD81;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Epidermis;
 RX MEDLINE=96178080; PubMed=8647901; DOI=10.1083/jcb.132.4.727;
 RA Xia Y., Gil S.G., Carter W.G.;
 RT "Anchor age mediated by integrin alpha6beta4 to laminin 5 (epiligrin)
 RT regulates tyrosine phosphorylation of a membrane-associated 80-KD
 RT protein.";
 RL J. Cell Biol. 132:727-740 (1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Epidermis;
 RX PubMed=14739293; DOI=10.1074/jbc.M309678200;
 RA Brown T.A., Yang T.M., Zaitsevskaia T., Xia Y., Dunn C.A., Sigle R.O.,
 RA Knudsen B., Carter W.G.;
 RT "Adhesion or plaemin regulates tyrosine phosphorylation of a novel
 RT membrane glycoprotein p80/GP140/CUB domain-containing protein 1 in
 RT epithelia.";
 RL J. Biol. Chem. 279:14772-14783 (2004).
 DR EMBL; AY375452; AAR21289.1; -, mRNA.
 FT NON TER 691 691
 SQ SEQUENCE 691 AA; 76906 MW; 3B18AD27BDBF7338 CRC64;

Query Match 82.3%; Score 3603; DB 2; Length 691;
 Best Local Similarity 99.9%; Pred. No. 2.8e-257;
 Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGNCGSIALGLVLLGAARLPARGAEAFIALPRESNTIVLTKGPTLLAKCYVI 60
 DB 1 MAGNCGSIALGLVLLGAARLPARGAEAFIALPRESNTIVLTKGPTLLAKCYVI 60
 QY 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKIDMGSPCFEGVLOPSTSLPT 120
 DB 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKIDMGSPCFEGVLOPSTSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLELQSPRLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQSPRLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIRKQEGVKALHPWFHPRVNVSFSTIANSSIRKLCIISVFEGESATLMSANY 240
 DB 181 GTVSRIRKQEGVKALHPWFHPRVNVSFSTIANSSIRKLCIISVFEGESATLMSANY 240
 QY 241 PEGFPEDELMTQWVVPVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVFKLEDK 300
 DB 241 PEGFPEDELMTQWVVPVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVFKLEDK 300
 QY 301 QPGNAGFNLSLQGCDDQDQSPGILRLQFQVLVOHPONESNKIYVVDLSNERAMSLTIE 360
 DB 301 QPGNAGFNLSLQGCDDQDQSPGILRLQFQVLVOHPONESNKIYVVDLSNERAMSLTIE 360
 QY 361 PRPVQSRKFPVGCFCVLCIESRTCSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420
 DB 361 PRPVQSRKFPVGCFCVLCIESRTCSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420

Db 361 PRVVKQSRKRVPPCCFVCLSRITCSSNMLITGSSKAKISFELCDDLRLMNMNVEKTIISCTDH 4200

QY 421 RYCQKRSYSIQVPSDILHLPEVLEHPSMKLILPKDRLSLVLPAPKIQOHTHEKPCNTSF 4800

Db 421 RYCQKRSYSIQVPSDILHLPEVLEHPSMKLILPKDRLSLVLPAPKIQOHTHEKPCNTSF 4800

QY 481 SYLVASALIPSDILIFSSFCPPGGSIKIQYKQNIISTLTLPAPSPFOEASRQGLTVSIPY 5400

Db 481 SYLVASALIPSDILIFSSFCPPGGSIKIQYKQNIISTLTLPAPSPFOEASRQGLTVSIPY 5400

QY 541 FKEEGVFTVPPDTSKSYLLRTPMWRGRLSLTSVMNLSVPPDQVACLTFFKERSGVVQ 6000

Db 541 FKEEGVFTVPPDTSKSYLLRTPMWRGRLSLTSVMNLSVPPDQVACLTFFKERSGVVQ 6000

QY 601 TGRAFMILIOEORTRAEILFSLDEDAVLPKPSFHHSFWNISNCSPTSGKQDILLFSVTLT 6600

Db 601 TGRAFMILIOEORTRAEILFSLDEDAVLPKPSFHHSFWNISNCSPTSGKQDILLFSVTLT 6600

QY 661 PRYDVLVILIAAVGGVLLLSLGLITTCV 691

Db 661 PRYDVLVILIAAVGGVLLLSLGLITTCV 691

Query Match	81.7%	Score 3574.5	DB 2	Length 833
RESULT 5				
OS0462_MOUSE		PRT	833 AA.	
ID OS0462_MOUSE PRELIMINARY;				
AC OS0462;				
DT 01-FEB-2005 (TrEMBLrel. 29, Created)				
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)				
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)				
DE CUB domain-containing protein 1.				
GN Name=Cdcp1;				
OS Mus musculus (Mouse).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;				
OC Muridae; Murinae; Mus.				
OX NCBI_TaxID=10090;				
[1]				
NP NUCLEOTIDE SEQUENCE.				
RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;				
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge U.G.,				
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,				
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,				
RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Heien F.,				
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schreetz T.E.,				
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,				
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,				
RA Bosak S.A., McEwan P.J., McKernan K.U., Malek J.A., Gunnarone P.H.,				
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA Fahey J., Helton E., Kettunen M., Madau A., Rodrigues S., Sanchez A.,				
RA Whiting M., Madau A., Young A.C., Shevchenko Y., Boulfard G.G.,				
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA Buterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalace D.E.,				
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,				
RT "Generation and initial analysis of more than 15,000 full-length human				
RT and mouse cDNA sequences."				
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
RL [2]				
RP NUCLEOTIDE SEQUENCE.				
RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;				
RG NIH WGC Project;				
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.				
DR EMBL: BC085253; AAI85253.1; -, mRNA.				
DR GO: GO:0016021; C: integral to membrane; IEA.				
DR InterPro: IPR000859; CUB.				
SO SEQUENCE 833 AA; 92725 MW; 1F142D3D595C6B27 CRC64;				

Best Local Similarity 80.9%; Pred. No. 4,7e-255;
Matches 676; Conservative 60; Mismatches 97; Indels 3; Gaps 3

Oy	1	MAGANCVSALLGVLLIGARLRGRGAEEFIALPRESNIIVLTKLGPTTLAKCYIV	60
Db	1	MAHSACGSVALLDALLLGTARLRLGTAEIALPQRGSVVISLQGNALPVLKCYIVM	60
Oy	61	SKRHITMISISGERIVFTFSQSPENHFVLEIQKNIIDCMGSPCFGSEVOLQPSLPIPT	120
Db	61	SRQITELLIRPGRKSFTECSNPEKHFVLAKENIDCMGSPCFGSEVHLQPSISELPI	120
Oy	121	LNRFPIWDVKAHKSITGLEQFSIRPLROIIGGESCPDGVTHSISGRIDATVIRISTFCSN	180
Db	121	LNRFPIWDVRAHKSITGLEQFATRLROIIGGESCADGVTHSISGHIDATEVIRIGTFCSN	180
Oy	181	GTVARIKMOEVSVMALHLPWFHPNNVSQFSIANSSTIRLCITFSVEFEGSATALMSANY	240
Db	181	GTVARIKMOEVSVMKALHLPWFHRNNVSQFSIANSSTIRLCITFSVEFEGSATALMSANY	240
Oy	241	PEGPEPEELMTQCVFVPAHLRPAVSFLNFNLSNCRKEERKEYIIPGSTTPEVFKLBDK	300
Db	241	PGGPEPEELMTQCVFVPAHLRPAVSFLNFNLSNCRKEERKEYIIPGSTTPEVFRLEBK	300
Oy	301	QPGNMAGNFNLSLQCCDDAQSPEGIRLQFQVLYQHPQESNKITYVDLSNERANSLTTE	360
Db	301	QPGNMAGNFNLSLQCCDDAQSPEILRLQFQVLYQHPQESNKITYVDLSREERTMSLTTE	360
Oy	361	PRPVKQARKPVPGGFVCLSESTCSNLITSGSHKISFLCDDLIRLMMANKEKITSCTDH	420
Db	361	PRPVKHGRFPVPGGFVCLSESTCSNTVYLTIGSHKISFLCDDLIRLMMANKEKITSCLDH	420
Oy	421	RYCQKRSYSIQVPEDILHLPELHDPFWKLLVPPDRSLVLVPAQKLQOHTHEKPCNTSF	480
Db	421	RYCVRQSFVKIQVPYIIQLPVQLHDFSWKLLVPPDKLSMLVPQKLIQHTHERCNTSF	480
Oy	481	SYLVASAIPSODLYFGSFCPEGSIKQIQVKNONISVTLRTFPAFSPFOEASRQGLYSLFIPY	540
Db	481	GYHTSSTTPGODLIFYGFSFGSGSIEKIQVKNONSVTTLRAVAPSPFOEVSXQGLIYVSYIPY	540
Oy	541	FKEGSGVTLVPEDTSSXYLRTPPMNDRGPLSTLTSWMSISVPRDOVALCTFPERSGVYVQ	600
Db	541	FKEGSGITLVVPDITKXKYLRPPMNDRGPLALSSVSWNISVSNDVALCTLVLEKSGLACQ	600
Oy	601	TGRAFMIIQEQRTAAEEIFSLDEQVLPKPSFHHNSFWNINSGSPSGSKOULLFSVTLT	660
Db	601	SGRAFMIIQEQQSGAAEEIFSLDEQVLPKPSFHHNSFWNINSGSPMGKODULLFSVTLT	660
Oy	661	PRIVDLTVILIAAVGGVLLISALGLITCYVKKKKKKTKTKGPAVGITNKNINTEKPROPK	720
Db	661	PRIVDLVAV-IGAAGGALLLFAVLITICF-KKKKKYDKKPAVGIVGNVNTQMP-OTQ	717
Oy	721	KFOKGRDNDSHVAVIEDTMVYGHLLQDSSGSLQEPVDVYRPPQGMVGCPEPPLTIC	780
Db	718	KFPKGRDNDSHVAVIEDTMVYGHLLQDSSGSLQEPVDVYRPPQGMGOCPPPPPLF	777
Oy	781	SRAPTAKLATEEPPRSPSESESPYFHSNNNDVASKDIDPLLTKQPEMEPAE	836
Db	778	SRTTAKFTABELALPSPSESESEPYFHSNNKEIGVRENDIDPLHTQGVETE	833
RESULT 6			
O9H8C2 HUMAN			
ID	O9H8C2	HUMAN PRELIMINARY;	PRT; 649 AA.
AC	O9H8C2		
DT	01-MAR-2001	(TReMBLrel. 16, Created)	
DT	01-MAR-2001	(TReMBLrel. 16, Last sequence update)	
DT	01-OCT-2002	(TReMBLrel. 22, Last annotation update)	
DE	Hypothetical protein FLJ13772.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;		
CC	Homo.		
NCBI	TaxID=9606;		

[1]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Placenta; DOI=10.1038/ng1285;
 EX PubMed:14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi K., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihira K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamakita M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Iosida M., Hotta T.,
 RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Mueshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Teraehima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Neguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsunura R., Kakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR GO: AK023834; BAB14695.1; -: mRNA.
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR InterPro: IPR000859; CDB.
 SQ SEQUENCE 649 AA; 72579 MW; 0F4404AD31D60E9A CRC64;
 Query Match 77.6%; Score 3395; DB 2; Length 649;
 Best Local Similarity 99.2%; Pred. No. 6.1e-242;
 Matches 644; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 188 MOEGVKAALHPHFPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANYPEGPED 247
 DB 1 MOEGVKAALHPHFPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANYPEGPED 60
 QY 248 ELMWQFVPAHLRASVSLNPNLSNCRKEERVVYTPGSTTNEVEVKLEDKOGNNAG 307
 DB 61 ELMWQFVPAHLRASVSLNPNLSNCRKEERVVYTPGSTTNEVEVKLEDKOGNNAG 120
 QY 308 NFNLISLQCCDDAOSPGLRLQFQVLYVQHPONESNKIYVVDLSNERAMSLTIEPRPV 367
 DB 121 NFNLISLQCCDDAOSPGLRLQFQVLYVQHPONESNKIYVVDLSNERAMSLTIEPRPV 180
 QY 368 RKEVPFGCVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVKITSCTDHRVCQKS 427
 DB 181 RKEVPFGCVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVKITSCTDHRVCQKS 240
 QY 428 YSLVQPSDILHLPVELHDFSWKLVLPKORLSLVLPVPAKLOOHTHEKCNISFSLVSA 487
 DB 241 YSLVQPSDILHLPVELHDFSWKLVLPKORLSLVLPVPAKLOOHTHEKCNISFSLVSA 300
 QY 488 IPSODLYFGSFCPGGSIKOIVKONISVTLRTFAPSFOEASROGLTVSFIPYKEEAVF 547
 DB 301 IPSODLYFGSFCPGGSIKOIVKONISVTLRTFAPSFOEASROGLTVSFIPYKEEAVF 360
 QY 548 TVTPTDTSKVKYLRTPNMVRGRLPSLTSVSWNISVPRDQVACLTFPERSGVVCQGRAPMI 607
 DB 361 TVTPTDTSKVKYLRTPNMVRGRLPSLTSVSWNISVPRDQVACLTFPERSGVVCQGRAPMI 420
 QY 608 IOEORTRAEBIFSLDEDEVLPKPSFHHHSFWNINSCSPTSKQJDLPSVTLTPTVDT 667

DB 421 IOEORTRAEBIFSLDEDEVLPKPSFHHHSFWNINSCSPTSKQJDLPSVTLTPTVDT 480
 QY 668 VILIAAVGGVLLSALGLIICVKKKKKKTKNGPAGVIYNNXINTENPRQPKFKQGRK 727
 DB 481 VILIAAVGGVLLSALGLIICVKKKKKKTKNGPAGVIYNNXINTENPRQPKFKQGRK 540
 QY 728 DNDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPPQGTNGVCPSPPTCSRAPYAK 787
 DB 541 DNDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPPQGTNGVCPSPPTCSRAPYAK 600
 QY 788 LATEPRSPSESESEPTSHPNNGVSSKDTDIPILTXQEPMEPAE 836
 DB 601 LATEPRSPSESESEPTSHPNNGVSSKDTDIPILTXQEPMEPAE 649
 RESULT 7
 ID 09H676 HUMAN PRELIMINARY; PRT; 392 AA.
 AC 09H676;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ22534.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
 OC Homo.
 ON NCBI_Taxid=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Ohyashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.,
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK026187; BAB15388.1; -: mRNA.
 FT NON TER 392 392
 SQ SEQUENCE 392 AA; 44030 MW; 3E21CBB2703B52F3 CRC64;
 Query Match 46.6%; Score 2041; DB 2; Length 392;
 Best Local Similarity 99.5%; Pred. No. 3.6e-142;
 Matches 390; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 305 MAGNPNLSLQCCDDAOSPGLRLQFQVLYVQHPONESNKIYVVDLSNERAMSLTIEPRPV 364
 DB 1 MAGNPNLSLQCCDDAOSPGLRLQFQVLYVQHPONESNKIYVVDLSNERAMSLTIEPRPV 60
 QY 365 KOSRKFVPGCVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVKITSCTDHRVCQ 424
 DB 61 KOSRKFVPGCVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVKITSCTDHRVCQ 120
 QY 425 RKSYSLOVPSDILHLPVELHDFSWKLVLPKORLSLVLPVPAKLOOHTHEKCNISFSLV 484
 DB 121 RKSYSLOVPSDILHLPVELHDFSWKLVLPKORLSLVLPVPAKLOOHTHEKCNISFSLV 180
 QY 485 ASAIPSDLYFGSFCPGGSIKOIVKONISVTLRTFAPSFOEASROGLTVSFIPYKEE 544
 DB 181 ASAIPSDLYFGSFCPGGSIKOIVKONISVTLRTFAPSFOEASROGLTVSFIPYKEE 240
 QY 545 GVFTVPTDTSKVKYLRTPNMVRGRLPSLTSVSWNISVPRDQVACLTFPERSGVVCQGRAP 604
 DB 241 GVFTVPTDTSKVKYLRTPNMVRGRLPSLTSVSWNISVPRDQVACLTFPERSGVVCQGRAP 300
 QY 605 FMIIOEORTRAEBIFSLDEDEVLPKPSFHHHSFWNINSCSPTSKQJDLPSVTLTPTVDT 664
 DB 301 FMIIOEORTRAEBIFSLDEDEVLPKPSFHHHSFWNINSCSPTSKQJDLPSVTLTPTVDT 360
 QY 665 DLTVILIAAVGGVLLSALGLIICVKKKKK 696
 DB 361 DLTVILIAAVGGVLLSALGLIICVKKKKK 392
 RESULT 8

Ox	NCB1_TaxID=9606;
RN	(1)
RN	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Kidney;
RX	MEDLINE=2228857; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strunberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA	Altshul S.F., Zeeberg B., Bucoff K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max I., Wang J., Heien F.,
RA	Dickchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Frange C.,
RA	Rana S.S., Loquellano N.A., Peters S., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fabry J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Tuchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez V.S.N., Grifwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA	Scherer A., Schein J.E., Jones S.J.M., Skalska U., Smalhus D.E.,
RT	"Generation and initial analysis of more than 15,000 full-length human
RL	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	(2)
RN	NUCLEOTIDE SEQUENCE.
RP	TISSUE=Kidney;
RG	NIH MGC Project:
RL	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN	(3)
RN	NUCLEOTIDE SEQUENCE.
RX	MEDLINE=2268796; PubMed=12975309; DOI=10.1101/gr.1293003;
RA	Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA	Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA	Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Helens S.,
RA	Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA	Lewis L., Liao D., Mark M.R., Robble E., Sanchez C., Schoenfeld J.,
RA	Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA	Vanden R.L., Watanabe C., Wiend D., Woods K., Xie M.-H.,
RA	Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA	Wood W.I., Godowski P.U., Gray A.M.;
RT	"The secreted protein discovery initiative (SPDI), a large-scale
RT	effort to identify novel human secreted and transmembrane proteins: a
RL	bioinformatics assessment."
RL	Genome Res. 13:2265-2270(2003).
DR	EMBL; BC021099; AAH21099.1; -; mRNA.
DR	EMBL; AY358779; AAC89139.1; -; mRNA.
DR	Ensembl; ENSG00000163814; Homo sapiens.
SO	SEQUENCE 343 AA; 37818 MW; 3BAEB13379DD94D1B CRC64;
<hr/>	
Query Match	40.6%; Score 1779; DB 2; Length 343;
Best local Similarity	99.4%; Pred.No.6,9e-123;
Match 341; Conservative	2; Mismatches 0; Indels 0; Gaps 0
Oy	1 MAGLNCGVSIALLGLVLLGAARLPKGAEEAFIALPRESNTIVLIKGTPTLAKPCYIY 60
Db	1 MAGLNCGVSIALLGLVLLGAARLPKGAEEAFIALPRESNTIVLIKGTPTLAKPCYIY 60
Oy	61 SKRHITMLSIKSGERIVFTFSCQSPBNHFVIEIKXNDICMSGCPGEVOLQPSLTAPT 120
Db	61 SKRHITMLSIKSGERIVFTFSCQSPBNHFVIEIKXNDICMSGCPGEVOLQPSLTAPT 120
Oy	121 LNRTIMVVKAHKSIGLEIOFSIPRLROIQBPESCDGYTHSISGRINDATVVRIRIGFCFN 180
Db	121 LNRTIMVVKAHKSIGLEIOFSIPRLROIQBPESCDGYTHSISGRINDATVVRIRIGFCFN 180
Oy	181 GTVSRIKIQEGVGMALHLPMFPAPRVNGSGSINRSSIKRLCTIESVFEEGSGATLMASNY 240
Db	181 GTVSRIKIQEGVGMALHLPMFPAPRVNGSGSINRSSIKRLCTIESVFEEGSGATLMASNY 240
Oy	241 PEGRPEELMTWQPVVPAPHLRASVSEFLNLNSCNCKEKREVERYYIPGSTTNPDEVFKLEDK 300
Db	241 PEGRPEELMTWQPVVPAPHLRASVSEFLNLNSCNCKEKREVERYYIPGSTTNPDEVFKLEDK 300

Db	241	PEGPEDELMTQWVFPVPAHLRASVSLNPNLSNCRKEREVEYIPGSIITPNPEVKLEDK	300
Qy	301	QPGNAGNFNLSTGCDODADQSPGILRIQFVLVGHQPONESNK	343
Db	301	QPGNAGNFNLSTGCDODADQSPGILRIQFVLVGHQPONESSE	343
RESULT 10			
Q921M9_MOUSE			
Q921M9_MOUSE	PRELIMINARY;	PRT;	384 AA.
AC	Q921M9_	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)	
DE	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Cdcpl protein (Fragment).		
GN	Name=E030027H19Rik; Synonyms=Cdcpl;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muroidea; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;		
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Klausner R. L., Fellings E. A., Grouse L. H., Derge J. G.,		
RA	Strausberg R. D., Collins F. S., Wagner L., Shennan C. M., Schuler G. D.,		
RA	Altschul S. F., Zeeberg B., Buetow K. H., Scheffer C. F., Bhat N. K.,		
RA	Hopkins R. F., Jordan H., Moore T., Max S. I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marzina K., Farmer A. A., Rubin G. M., Hong L.,		
RA	Stapleton M., Soares M. B., Bonaldi M. F., Casavant T. L., Scheetz T. E.,		
RA	Brownstein M. J., Uesdi T. B., Toshiyuki S., Caranci P., Prange C.,		
RA	Raba S. S., Loquellano N. A., Peters G. J., Abramson R. D., Mullaly S. J.,		
RA	Boesk S. A., McEwan P. J., McKernan K. J., Malek J. A., Gunaratne P. H.,		
RA	Richards S., Worley K. C., Hale S., Garcia A. M., Gay L. J., Huijck S. W.,		
RA	Villation D. K., Muzny D. M., Sodergren E. J., Lu X., Gibbs R. A.,		
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting R. W., Madan A., Young A. C., Shevchenko Y., Bouffard G. G.,		
RA	Blakeley A. C., Grimwood J., Schmutz J., Myers R. M.,		
RA	Rodriguez Y. S. N., Krzywinski M. I., Skalska J., Smalhus D. E.,		
RA	Schmerch A., Schein J. E., Jones S. J. M., Marra M. A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RL	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;		
RC	NIH MGC Project;		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; BC0111340; AAH11340.1; mRNA.		
DR	Ensembl; ENSMUSG00000035496; Mus musculus.		
DR	MGI; MGI:2442010; Cdcpl.		
DR	MGI; MGI:2442010; E030027H19Rik.		
FT	NON TER		
FT	1		
FT	1		
SEQ	SEQUENCE	384 AA; 42406 MW; 4B5F17D30FBE16080 CRC64;	
Qy	Query Match	36.1%; Score 1579.5; DB 2; Length 384;	
Qy	Best Local Similarity	77.3%; Pred. No. 4.6e-108;	
Qy	Matches 299; Conservative 34; Mismatches 51; Indels 3; Gaps 3		
Db	450	LLVPRKRLSLVLAPOAKLQOHTHEKPCNTSSYLAASAI PSODLYFGSFCPCGSGIKOIY	509
Db	1	LLVPRKRLSLVLAPOAKLQOHTHEKPCNTSSYLAASAI PSODLYFGSFCPCGSGIKOIY	60
Qy	510	KONISVTLRTAPSPFOEASRQGLTVSFI PFYKEGVETVTPDTKSKYVLRTPNMDRGLP	569
Db	61	KONISVTLRTAPSPFOEASRQGLTVSFI PFYKEGVETVTPDTKSKYVLRTPNMDRGLP	120
Qy	570	SLTSSVNNISVPRPOVACLTFPKERSGVVCTGGAFFIIQSKRTAEIIFSLDEVLKXP	629
Db	121	ALSSVNNISVPRPOVACLTFPKERSGVVCTGGAFFIIQSKRTAEIIFSLDEVLKXP	180

```

Oy 630 SFHHHSFWWNISNCSPTSGKOLDLFSYTLTPRTVDLTVILIAVGGVLLSALGLIIC 689
Db 181 SFHHHSFWWNISNCSPTSGKOLDLFSYTLTPRTVDLTVILIAVGGVLLSALGLIIC 239
Oy 690 CVKKKKKKTNNKPAVAGIYNNXINTEMPROPKKFGKRDNDSHVAVIEDTMYGHLLOD 749
Db 240 FV-KKKKKVKGSPAVGIYNNVNTQMP-QTKFPPKGRKDNDSHVAVIEDTMYGHLLOD 297
Oy 750 SSGSLQLEVDVITTRPFGTGMVCPSPSPITICSRAPAKLATEPPEPPSPSESEPTYES 809
Db 298 SGGSLQLEVDVITTRPFGTGMVCPSPSPITICSRAPAKLATEPPEPPSPSESEPTYES 357
Oy 810 HPNNNGDVSSKDTDIPDLXTOEPMPEPAE 836
Db 358 HPNNNGDVSSKDTDIPDLXTOEPMPEPAE 384

RESULT 11
OAT223 TETNG PRELIMINARY; PRT; 867 AA.
ID OAT223 TETNG PRELIMINARY; PRT; 867 AA.
AC OAT223 TETNG PRELIMINARY; PRT; 867 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF10362, whole genome shotgun sequence.
GN ORFNames=GSTENG0000853001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vachette B.,
RA Blemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., McEwan P., Bosak S.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., Zody M.C.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mestrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Querier F., Saurin W., Scarpelli C.,
RA Winkler P., Landet E.S., Weissenbach J., Roest Crollius H.,
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RA the early vertebrate proto-karyotype."
RA Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Genoscope, Whitehead Institute Centre for Genome Research;
RA Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAAB01010362; CAP93059.1; -: Genomic DNA.
SQ SEQUENCE 867 AA; 96120 MM; 1177A6DEBABA88A CRC64;

Query Match 18.9%; Score 828; DB 2; Length 867;
Best Local Similarity 30.0%; Pred. No. 5.4e-52;
Matches 243; Conservative 135; Mismatches 311; Indels 120; Gaps 32;

```

```

Db 220 VLYKALMTLQVGHRELDPVDFTLNSNGPSTMTAVKVSILPRGVSTGFFAANYPNDEPD 279
Oy 247 DELMTWQVNVFAHLBASYSFLNPNLSNCRKEEVEYIIPGSTNPEVFKLEDKQPGMA 306
Db 280 KOQVEWDFVGMENYVTHFAHTAPBCLNGKEVEEENKQKQSVTR--LALTDPOPEHQ 337
Oy 307 GNPNLSLQGCQDQASPGILRLQOVLVQ--HPONESNKIYVLDLSNERAMSLTIEBRP 363
Db 338 GDFKAVLKNKGTNTNLQG-LFLDYKVSVMBSGHP-----VLCTYDLTGKQVSLQD----- 387
Oy 364 VKQSRKFEVGGFVCLIESFTCSNLTLSGSKHKSIFLCDLDRILMNVKRTISCTDHYXC 423
Db 388 LKVGSD--PYCEMSVNSK-IEKINVAAGTILASISFLDCPSQDVLRTATKTIACQHLGSC 444
Oy 424 QKSYSLQVPSDILHLPVLEHDFSWKLVLPKDRSLVLPVPAQKLOQHTHERPCNTSPSYL 483
Db 445 --PSSLTVPSLDSCLPLPHSFTMHMMIIPDSTYDMLSPGSLHQSLPQGCBSLSLH 502
Oy 484 VASAIPSDLYFGSFCPGSIKQIQVKONISVTLRTFAPSFXQEASRQGLTVSFLPYFKE 543
Db 503 VAE--SDGLSLGDFCFNGAIOKIOAHNISTVAR--VPDF-KKSRGPFLLASFVQEIQD 556
Oy 544 EGFTVTPDYTSKYVLRTPNMDRGLPSLTSVSNISVPRDOVALCTFKKESGVVC--Q 600
Db 557 TFIYRISPE-EPQTLATPMPMGKPSSTVSWVTLPSQYEAADRFFNV-SQPKDNH 614
Oy 601 TGRAMIIQEQRTAREIPLSDED-----VLPRKSFHHHSFWWNISNCSPTSGKOLDL 653
Db 615 TSIKWMLQGE--DELSRREDPKYDMLVPE-----SFIYAMNCLIPQEG-QRKA 663
Oy 654 LFSYTLTPRTVDLTVILIAVGGVLLSALGLIICVKKKK--KTNKPAVGIYNNXI 711
Db 664 MAKIVLQKSNLPLIIL--GVAAGAFILLLIILAVVCVKKKKSKKANKNSIYWKAKSI 721
Oy 712 NTEMPROKPKQ-KCRKNDSHVAVIEDTMYGHLLODS-----GSFLQPEVDYTRP 764
Db 722 ----FRPDEIRFKSRSDNNHIYDSIDETWYGHLPDSDLSLMSYSQMVDSTYOT 776
Oy 765 FQG-TMGVCPSPSPITICSRAPAKLATEP-----PPRS--PPES----- 801
Db 777 FSGPTDGKLP-----VIEPDHDEVEQGTFLDPKSPMPRPPTPIDR 821
Oy 802 -----ESEPYTFSHPNNDVSS 818
Db 822 QDSLGFQDSRMVNDNLYTFK--STGDMYT 848

RESULT 12
OAT223 TETNG PRELIMINARY; PRT; 825 AA.
ID OAT223 TETNG PRELIMINARY; PRT; 825 AA.
AC OAT223 TETNG PRELIMINARY; PRT; 825 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF7985, whole genome shotgun sequence.
GN ORFNames=GSTENG0000554001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Catolico L., Poulain J., Katinka M., Vachette B.,
RA Blemont C., Skalli Z., Catolico L., Coutanceau J.P., Guzy J.,
RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., McEwan P., Bosak S.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., Zody M.C., Mestrov J.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mestrov J.,

```


DT 10-MAY-2005 (TrEMBLrel. 30, last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, last annotation update)
 DE Hypothetical LOC541520.
 GN Name=zgc:113071;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_Taxid=7955;
 OK [1]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Singapore local strain; TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schejter C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield J.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Singapore local strain; TISSUE=Embryo;
 RC NIH MGC Project;
 RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 RL EMBL: BC091838; AAH91838.1; -, mRNA.
 KM Hypothetical protein.
 SQ SEQUENCE 793 AA; 88434 MW; C941FC39F0F9F3A9 CRC64;
 Query Match 13.9%; Score 610; DB 2; Length 793;
 Query Local Similarity 25.7%; Pred. No. 6,1e-36;
 Matches 215; Conservative 153; Mismatches 372; Indels 98; Gaps 30;
 QY 9 STALLGVTLLGAAALPRGAFAEIALPRESNTVILIKGPT-LLAKPC----- 56
 DB 10 SLVLLVFLSSVLELTGAE-LINVA--EQGTILINTSSSTGCKVMRYANQPLQA 66
 QY 57 YIVISKRTITMSTISGERIVTFSCQSPENHFVIEICKIDCMSPGCPFGVQLQPS- 115
 DB 67 YVVL--RSPTLDD-----FSCSQPEKVFIVQIDT-----GTMNLEPKCH 105
 QY 116 SLPLTLNRFIVDVKAHKSIGLEQPSIRLQIGGSCGPDGVTISGRIDATVVRIG 175
 DB 106 SAPIKRNRFITNKKLVSKSLQIANSSSTGIQIIEPDSCPNHRKIYA-----VENISIG 161
 QY 176 TFCNSGTSGRIKMGQGVKALHLPMWPRNVSFGSIANSSIKRLCIISVEGEGSATL 235
 DB 162 TFCQGVIRIQIDVRNAGKLSAEVTGGQSISTKVIASVGRMLNSLVNIHAI,PEKSS--I 219
 QY 236 MEANYPGEPFEDLMTWQVVPRAHLRAYSFLNPNLSNCKEKE--RVETIYIPGSTTNE 293
 DB 220 QDFITPNTTPENAETWFMFSPMAVYADRIILKYTLPTCISHSISPTMKYTWG--RDAL 277
 QY 294 VFKELEQKQGNMAGNFNLSLOGDD--AOSPGLRLQFOFV-LYQHPONSNNKYVVD,SN 351
 DB 278 VGRMENOQSVBEGANNISIKCKMSTVSLNRLMWHFQISAIKRRKQCEG----DLPD 333
 QY 352 ERAMSLTIPRPVKQSRKVPKPGCVLESRTSSNLTLSGSKHKSIFLCDDLTRLMNV 411
 DB 334 KQVLKIQV---MKQPKSV--CVLKLDS-VIMDTVTIASGNFIINTFDCKNDELRLV 386

QY 412 EKTISTDHRXCQKRSYSIQVPSDILHPEVLEHDFSMKLVKDRSLVLPAGKIQHT 471
 DB 387 IQTECKKRWKCTSTLPFLHRYNEQCCIPGLKMTWTYIQGPVNAVVLQSTDLARCL 446
 QY 472 HEKPCNTSPSYLVASAPSDLYFGSFCPGSGIKQIQVKON-ISVTLTFAPSFYQASR 530
 DB 447 PEDNCC-SILLNVSQVNNPGTWQFCPKGSVQKIQRESKIATVASTSPSDSVAAK 505
 QY 531 QGLTVSFPYKEBEVFTVTPDTISKVYLRTPNMVRGLPSLTSVSMNLSVPRDQVACLF 590
 DB 506 PILSYSTFOGISSENIFFSVTPMDPPTLAPPGWPSAKASTSVIWMPEQYTSVLT- 564
 QY 591 FKERSGVCCOTGRAPMIIOEORTAEIIFSLDEVLPRKSFHHSFWMNISNC-SPTSGK 649
 DB 555 FRANSQPKCKQVHTNIAVQITRSQ-VLYTTKDKMDLLPBEFYLMNTCKSPGTA- 622
 QY 650 QLDLFSVTLTPRTVDLTVIILAAVGGVLLLSALGLIICVKKKKKTKNGPAVGIIYNX 709
 DB 623 -FRAMWEVTQSNTKLGIILGVV--GVVLAIAAVAVIWFVLRKKR-NKAPVSVYNP 678
 QY 710 NINTMPROPKFKQGRKDNDSHYAVIEDTMYGHLQDSSGSLQPEVDTRYRFGQTM 769
 DB 679 SEHAPLPGIHGIPKQGEEDENRTYAVYIDTLVYSHLND-----DADKKQHYEDTS 730
 QY 770 GVCPESP-----TICSRAPTAKL-----ATEEPPRSPSESESEBYTF 808
 DB 731 GPHLAKPPLPRDPDNNVSHALDKLLPNELYGYVGQSTSGPRTKAGS-SEHTY 787
 RESULT 15
 001654 HALRO
 ID 001654 HALRO PRELIMINARY; PRT; 746 AA.
 AC 001654;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
 DE AsmSPa.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 NC NCB1_Taxid=7729;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Healopancreas;
 RC MEDLINE=97322374; PubMed=9177219; DOI=10.1073/pnas.94.12.6340;
 RA Uj X., Azumi K., Sasaki M., Nonaka M.;
 RT "Ancient origin of the complement lectin pathway revealed by molecular
 RT cloning of mannan binding protein-associated serine protease from a
 RT urochordata, the Japanese Ascidian, Halocynthia roretzi.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:6340-6345(1997).
 DB EMBL: D88204; BAA19762.1; -, mRNA.
 DR HSP; Q9JUS8; INTO.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006956; P:complement activation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001254; Peptidase_S1_S6.
 DR InterPro; IPR011361; Pept_S1_Comp_Act.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00089; Sushi; 1.
 DR Pfam; PF00089; Trypsin; 1.
 DR PIRSF; PIRSF001155; Clt_Cig_MASP; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00032; CCP; 2.
 DR SMART; SM00042; CUB; 2.

DR SMART; SM00179; BGF_CA; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00010; ASK_HYDROXYL; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50923; SUSHI; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 746 AA; 84127 MW; 32B075317CF173E1 CRC64;

Query Match 3.2%; Score 141; DB 2; Length 746;
Best Local Similarity 19.3%; Pred. No. 0.26;

Matches 102; Conservative 73; Mismatches 178; Indels 176; Gaps 25;

QY 220 LCITISVEEGEGSATLMSA-----NYPEGFPEDELMTQFVVPAAHLRASVSFLNENL 271
DB 11 LLYCNSVF---SAILLTAHFGNFPSSPNYPSPYDPSNLTWNIRVQHGYRMSIRPSTFDL 66
QY 272 -----SNCRKKEKREVEYIRGSTTNEVEFKLEDKOPGNMAGNPNLSLQCCDDAOSP 323
DB 67 EDSYEDGIGSC-----YDYVEITESNKTVAK-----FCGNVQL---FPTDAPNP 108
QY 324 GILRLQFQVLVQHPQNESNKIYVVDLS-----NERAMSLTIEPPRPVKOSRK----- 369
DB 109 -----SKFITSGNEVRVTFVSDYSLSGFOAHYAQIDINECELMETKRTIIEDMD 161
QY 370 -----FVPGCFVCLIESRTGSSNLTLSGSRHKISFLCDDLTRLMMNVEKTICTD 419
DB 162 ELVVCSHYCRNVPGSYYC---GCRPKFTLDNRHTCVASFCE-----QVLTDDN 208
QY 420 HRYCCOKSKYSLOVPDILHLPELHDFGSKLVPRD-RLSLVLVPAQKLQCHTHEKPCNT 478
DB 209 SGHISSEPP-----ELVAKLTDCSWTIQLREGLSVNLIFERAFGIEHEEE--GC 257
QY 479 SFSYLVASAIPSODLYFGSFCPG-----SIKOIVKONISYTLRTFAPSPYQ--- 526
DB 258 TYDRLAEVLHKSTTDIYCGNQAPNGTVMPMNTNLVQLKHTDLSYEKKGFSAVRYTSTRIK 317
QY 527 -----EASRQGLTVSFI---PYKKEGVFTVPTDTSKVVYLRTPNMDRGLPSLTSV---S 575
DB 318 CLHALHDPKNG-SLSFSSHRSYHEFEDVATFSCDRGFDII-----GVPRKLCLSDGS 368
QY 576 WNISVPRDQV-AC-----LTFEKRSGVQC-----TGRAFMII 608
DB 369 WSHSAPICQIKSCGVPQFLDLPNSHIVEYENSKTYSBVLDTVCNQWYGMISGASKWIC 428
QY 609 QEORTRAEEIFSLDEDLVLPKPSFHHSFPVNISN-----CAPTSGKOL 651
DB 429 ENSKIWTE-----HGGIVAINNFNNKPVCKPICGKTL 460

Search completed: May 4, 2006, 22:15:05
Job time : 125 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:01:05 ; Search time 49 Seconds
(without alignments)
1410.548 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4392
Sequence: 1 MAGUNCGVSTALLGVLLGA.....SSKDTDIPILSTQPMPEAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

No matches found

Search completed: May 4, 2006, 21:02:36
Job time : 49 secs

THIS PAGE LEFT BLANK